

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:48:34 ; Search time 20 Seconds
(without alignments)
1211.725 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 1348
Sequence: 1 MAELEFVQIIIIIVVMVMV.....PLESRAIWSKEKDQKGHPL 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125.5	9.3	634	2 T00388	hypothetical prote
2	123.5	9.2	508	2 T09046	proline-rich prote
3	122	9.1	530	2 T48627	hypothetical prote
4	113	8.4	628	2 S19150	hypothetical prote
5	111.5	8.3	1184	2 G01763	atrophin-1 - human
6	111.5	8.3	1184	2 S50832	atrophin-1 - human
7	111	8.2	727	2 C84534	hypothetical prote
8	110.5	8.2	778	2 I38487	tastin - human
9	109.5	8.1	1099	2 A56155	tumor suppressor p
10	107	7.9	2357	2 A59249	class VII unconven
11	105	7.8	1145	2 T33606	hypothetical prote
12	103	7.6	628	2 S01955	hypothetical prote
13	102.5	7.6	1357	2 T29265	hypothetical prote
14	101	7.5	384	2 S51796	vasodilator-stimul
15	101	7.5	944	2 T28734	hypothetical prote
16	101	7.5	1571	2 T00062	hypothetical prote
17	100.5	7.5	760	2 T06291	extensin homolog T
18	100	7.4	363	2 T16755	hypothetical prote
19	100	7.4	445	2 S00256	Krox-20 protein -
20	100	7.4	628	2 JQ0110	hypothetical 69K p
21	100	7.4	1132	2 T49403	related to protein
22	99.5	7.4	786	2 T01456	extensin homolog F
23	99.5	7.4	872	2 S33015	hypothetical prote
24	99	7.3	448	2 T06076	proline-rich prote
25	99	7.3	571	2 T43456	hypothetical prote
26	99	7.3	1051	2 T51904	hypothetical prote
27	98.5	7.3	1366	2 B84924	hypothetical prote
28	98	7.3	761	2 C84672	hypothetical prote
29	98	7.3	1172	2 T00065	hypothetical prote

30	97	7.2	527	2 B70700	hypothetical prote
31	96	7.1	436	2 T15331	hypothetical prote
32	96	7.1	620	2 S06733	hydroxyproline-ric
33	96	7.1	710	2 D96728	hypothetical prote
34	96	7.1	895	2 C86371	99.7K hypothetical
35	96	7.1	1460	1 EDBEIF	immediate-early pr
36	96	7.1	1844	1 RRWPTM	genome polyprotein
37	95.5	7.1	707	2 A46302	PTB-associated spl
38	95.5	7.1	998	2 T30930	hypothetical prote
39	95.5	7.1	1240	2 JC5209	insulin receptor s
40	95.5	7.1	5762	2 A41819	proline-rich pepti
41	95	7.0	876	2 A49508	proteins-tyrosine k
42	95	7.0	896	1 A35782	cytokine receptor
43	95	7.0	951	2 T47617	extensin-like prot
44	94.5	7.0	310	1 PIHUSD	salivary proline-r
45	94.5	7.0	322	2 S25299	extensin precursor

ALIGNMENTS

RESULT 1

T00388
hypothetical protein KIAA0616 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000

C;Accession: T00388

R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.

DNA Res. 5, 169-176, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete

A;Reference number: Z14142; MIM: 98403880; PMID: 9734811

A;Accession: T00388

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-634 <ISH>

A;Cross-references: EMBL:AB014516; NID:g3327045; PIDN:BAA31591.1; PID:g3327046

A;Experimental source: brain

C;Genetics:

A;Gene: KIAA0616

Query Match 9.3%; Score 125.5; DB 2; Length 634;
Best Local Similarity 24.7%; Pred. NO. 0.026;
Matches 53; Conservative 24; Mismatches 63; Indels 75; Gaps 8;

Qy	38	ISRHSGRRRE-----DALSGCLWPSESTVSGNGIPEQVYAPRRP	80
Db	314	LSLSTEARQQASPTLSPLSPITQAVAMDALSLEQLPYAFFTQAGSQPPPPQPPPPP	373
Qy	81	TDLAVPPFAQRERFHRFQPTYPLQHEIDL-----PTISLSDGEEPPP--YQGPCVL	132
Db	374	-----PPASQ-----PPPPPPQAPVRLPPGGLPLPSASLTRGPQPPPLAVTPSSL	421
Qy	133	QLRDPEQ-----QLELNRESVRAPPNRTIFDSDLMDSARLGGCPSPSSNSG	178
Db	422	PQSPENPGQPSMGIDIASAPALQQYRTSAGSPANQS-----PTSPVSNQG	467
Qy	179	ISATCYGSGRMEGPPPTSEVIGHYPGSSFOHQ	213
Db	468	FS-----PGSSPQHTSTLGSVFGDAYVEEQ	492

RESULT 2

T09046
proline-rich protein F26K10.180 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999

C;Accession: T09046

R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro

submitted to the Protein Sequence Database, June 1999

A;Reference number: Z16533

A;Accession: T09046

A;Molecule type: DNA

A;Residues: 1-508 <BEV>

RESULT 6
S50832
atrophin-1 - human
C;Species: Homo sapiens (man)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 07-May-1999
C;Accession: S50832
R;Nagafuchi, S.; Yanagisawa, H.; Ohsaki, E.; Shirayama, T.; Tadokoro, K.; Inoue, T.; Yam
Nature Genet. 8, 177-181, 1994
A;Title: Structure and expression of the gene responsible for the triplet repeat disorder
A;Reference number: S50832; MUID:95144175; PMID:7842016
A;Accession: S50832
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1184 <NAG>
A;Cross-references: EMBL:D31840
C;Genetics:
A;Gene: GDB:DRPLA; B37
A;Cross-references: GDB:270336; OMIM:125370
A;Map position: 12p13.31-12p13.3112p-12p

Query Match 8.3%; Score 111.5; DB 2; Length 1184;
Best Local Similarity 24.7%; Pred. No. 0.68;
Matches 48; Conservative 14; Mismatches 67; Indels 65; Gaps 7;
QY 52 SSEGCLWPSESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDL 111
Db 380 SSSAAASSSSSSSSASFPASQA-----LPSYPH-----SF 412
QY 112 PPTISLSDGEEPPPYQGCPCTLQLRDPQQLNRESV--RAPNRTIFDSDLMSARLGG 169
Db 413 PPPTSLSVSNQPKYTQP-----SLPSQAVWSQGPPPPPYGRLLANSNAHPG 460
QY 170 PCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQ-----SSGP----- 217
Db 461 PFPES-----TGAQSTAHPVPVSTHHHHQOQQOQQOQQOQQHHGNSGPPPPPGA 509
QY 218 -PSLLEGTRLHHTH 230
Db 510 FPHPLEGGSSHHAH 523

RESULT 7
C84534
hypothetical protein At2g15880 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84534
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84534
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-727 <STO>
A;Cross-references: GB:AE002093; NID:G5306245; PIDN:AAD41978.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g15880
A;Map position: 2

Query Match 8.2%; Score 111; DB 2; Length 727;
Best Local Similarity 25.0%; Pred. No. 0.43;
Matches 40; Conservative 18; Mismatches 54; Indels 48; Gaps 7;
QY 70 PEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGP 129
Db 609 PPPPYSPPPPPVHSPPPV-----FSPPPP-----VHSP-----PPVYSP 647
QY 130 CTQLQRPDPEQQLNRESVRAPNRTIFDSDLMSARLGGCPPPSSNSGISATCYGS--- 186

Db 648 PPVYSPPP-----PPVKSPPPPPVYSPLL-----PPKMSSPPTQTPVNSPPP 690
QY 187 ---GGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPPSLLEG 223
Db 691 RTPSQTVEAPPSEEFIIPFFIG----HQYASPPPPPMFQG 726
RESULT 8
I38487
tastin - human
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 05-Nov-1999
C;Accession: I38487
R;Fukuda, M.N.; Sato, T.; Nakayama, J.; Klier, G.; Mikami, M.; Aoki, D.; Nozawa, S.
Genes Dev. 9, 1199-1210, 1995
A;Title: Trophinin and tastin, a novel cell adhesion molecule complex with potential inv
A;Reference number: I38487; MUID:95278733; PMID:7758945
A;Accession: I38487
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-778 <RES>
A;Cross-references: EMBL:U04810; NID:G905355; PIDN:AAA79333.1; PID:G905356
Query Match 8.2%; Score 110.5; DB 2; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.51;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;
QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQRERFHRFQPTY--PYLQHEIDLPTTI 115
Db 498 SGLPKPCLPEECGEPCPPAEPG-----PP-----EAFCSRSEIPEPSLQEQLEVPEPY 548
QY 116 SLSDGEEPPPYQGCPCTLQLRDP-----QQLELNRESVRAPP-----NRTIFDSDL 162
Db 549 PPA---EPRPLESCCRSEIPESSRQEQLEVPEPCPPAEPRLSYCRIEPIESSRQ 605
QY 163 DSARLGGPCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPGQLQSTQGSQ--PPGPCPRVE--LGASEPCTLEHRSLESSLPPCC 662
QY 221 LEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 663 -----SQWAPATTSLIFFSQ-----HPL 680

RESULT 9
A56155
tumor suppressor protein warts (EC 2.7.1.1-) - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 24-Sep-1999
C;Accession: A56155
R;Justice, R.W.; Zilian, O.; Woods, D.F.; Noll, M.; Bryant, P.J.
Genes Dev. 9, 534-546, 1995
A;Title: The Drosophila tumor suppressor gene warts encodes a homolog of human myotonic
A;Reference number: A56155; MUID:95212904; PMID:7698644
A;Accession: A56155
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1099 <JUS>
A;Cross-references: GB:L39837; NID:G755007; PIDN:AAA73959.1; PID:G755008
C;Genetics:
A;Gene: wts
A;Cross-references: FlyBase:FBgn0011739
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C;Keywords: ATP; phosphotransferase
F;711-1014/Domain: protein kinase homology <KIN>
F;719-727/Region: protein kinase ATP-binding motif
Query Match 8.1%; Score 109.5; DB 2; Length 1099;
Best Local Similarity 26.6%; Pred. No. 0.9;
Matches 62; Conservative 18; Mismatches 80; Indels 73; Gaps 14;
QY 28 SHYKLSARSFIRHSQGRREDALSSGCLWPSESTV---SGNG--IPEQVY---APPR 79

Db 140 SHYLRCSPALDS--GAGSSRSDSPHSHHTQPSRTVGNPGNGGFSPPSGFSEVAPPA 197

QY 80 PTDR-----LAVPPFAQRERFHRFQPTYPYLQHEI-----DLPTTISLSDGEEPPP 125

Db 198 PPRNPTASSAATPP-----PPVPPTSQAYVKRRSPALNRPPIA-----PPT 241

QY 126 YQG--PCTLQ--LRDPEQQLNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISA 181

Db 242 ORGNSPVITQNLKNPQQLTQQLKSLNLYPGG-----GSAVVEPPPPYLIQG--- 290

QY 182 TCYSGGRMEGPP-----PTYSEVIGH--YPGSSFHQHQSSGPPS 219

Db 291 ---GAGGAAPPPPPSYTASMQRSQSTQSQDYKSPSSGIYSATSAGSPS 340

RESULT 10

A59249

Class VII unconventional myosin - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000

C:Accession: A59249

R:Titus, M.A.; Kuspa, A.; Loomis, W.F.

Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994

A:Title: Discovery of myosin genes by physical mapping in Dictyostelium.

A:Reference number: Z20873; MUID:95023928; PMID:7937787

A:Accession: A59249

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2357 <TIT>

A:Cross-references: GB:L35321; NID:G6226760; PIDN:AAF06035.1; PID:G6226761

A:Experimental source: gene myoI; product class VII unconventional myosin

R:Titus, M.A.

Curr. Biol. 9, 1297-1303, 1999

A:Title: A class VII unconventional myosin is required for phagocytosis.

A:Reference number: A59249

A:Contents: annotation

C:Genetics:

A:Gene: myoI

C:Superfamily: myosin motor domain homology

F:16-676/Domain: myosin motor domain homology <MMO>

Query Match 7.9%; Score 107; DB 2; Length 2357;

Best Local Similarity 28.1%; Pred. No. 3.3;

Matches 43; Conservative 12; Mismatches 56; Indels 42; Gaps 7;

QY 69 IPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQG 128

Db 1680 VPPQ---PVHPVATLSP-----MSPTIPNITN--TPPPPSISDSMSPPQVG 1724

QY 129 PCTLQLRDPEQQLNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISATCYGSGG 188

Db 1725 ----MLPPP-----PPPSVMGSKPIEIPSLGIPPPPPSSSSSVNPSPIGS 1767

QY 189 RMEG---PPPTYSEVIGHYPGSSFHQHQSSGPP 218

Db 1768 PMMGIPPPPTISV-----HSLNSNGNSTPP 1793

RESULT 11

T33606

hypothetical protein F54C4.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T33606

R:Rohlfing, T.; Antoniou, B.

submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid F54C4.

A:Reference number: Z21376

A:Accession: T33606

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1145 <ROH>

A:Cross-references: EMBL:AF099916; PIDN:AAC68776.1; GSPDB:GN00021; CESP:F54C4.3

A:Experimental source: strain Bristol N2; clone F54C4

C:Genetics:

A:Gene: CESP:F54C4.3

A:Map position: 3

A:Introns: 57/3; 98/1; 167/3; 193/1; 264/3; 307/3; 357/2; 371/2; 432/1; 515/3; 564/2; 63

C:Superfamily: Caenorhabditis elegans hypothetical protein F54C4.3

Query Match 7.8%; Score 105; DB 2; Length 1145;

Best Local Similarity 21.6%; Pred. No. 2.1;

Matches 49; Conservative 25; Mismatches 69; Indels 84; Gaps 8;

QY 86 VPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGCTLQLRDPEQQL-ELN 144

Db 929 IPP--HRQLYQKLDMPVISHELPRSP-----SRRIRIPTRKIQELD 969

QY 145 RESVRAPPNRTIFDSDLMD SARLG-----PCPPSSNSGISATCYGSGRME----- 191

Db 970 QESVLPPPPRRRTTRSVVRKRDRFAELRDCPPEVMRAIAAHAIAYDGRKKEORPIVY 1029

QY 192 -----GPPPTYSEVIGHYPGSSFHQHQ----- 213

Db 1030 VPRQAKMAMSTPTAPTPEDEDYDGPFGHQMTTILQHROAKQEEEEQWSGEYDDELHQ 1089

QY 214 -----SSGPPSLLEGTR---LHH---THIAPLESAAIWSKEDKQ 247

Db 1090 EHPATSSPPPSLLPHTGPSEIVHHEEVHMLPEEADLMDDDDQ 1136

RESULT 12

S01955

hypothetical protein, 69K - turnip yellow mosaic virus

C:Species: turnip yellow mosaic virus, TYMV

C:Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-2000

C:Accession: S01955

R:Morch, M.D.; Boyer, J.C.; Haenni, A.L.

Nucleic Acids Res. 16, 6157-6173, 1988

A:Title: Overlapping open reading frames revealed by complete nucleotide sequencing of t

A:Reference number: S01955; MUID:88289359; PMID:3399388

A:Accession: S01955

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-628 <MOR>

A:Cross-references: EMBL:X07441; NID:G62222; PIDN:CAA30321.1; PID:G62223

A:Note: the authors translated the codon ACG for residue 459 as U

C:Superfamily: hydroxyproline-rich glycoprotein

Query Match 7.6%; Score 103; DB 2; Length 628;

Best Local Similarity 25.3%; Pred. No. 1.6;

Matches 58; Conservative 19; Mismatches 80; Indels 72; Gaps 11;

QY 40 RHSQRRREDALSGCLWPSESTVSGN-----IPE---PQVYAPRPTDRLAVPPFAQ 91

Db 394 RH---RRPYLLPNPPALPSIAYTSSRGKIHSLPKALPKGAPP-PPRRLPSP--AP 447

QY 92 RERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGCTLQLRDPEQQLN----- 144

Db 448 R-----PQLPLRDLGRTPGFPPTPKTPTRTPESRITASPTDIAPL 487

QY 145 -----RESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISATCYGSGRMEGPPPT 196

Db 488 DSDPVLVSRTEVHAPERTFMDPEALRSALASLPSPRS-VGIHTA-----PQT 536

QY 197 YSEVIGHYPGSSFHQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKD 245

Db 537 ---VLPANPPSPTRHLPPPTSPWPILQ-----SPVGEDAIVDSEDD 573

RESULT 13

T29265

hypothetical protein C01G8.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29265

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:49:24 ; Search time 21 Seconds
(without alignments)
507.730 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 1348
Sequence: 1 MAELEFVQIIIIIVVMVMV.....PLESAATWSKEKDQKGHPL 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1348	100.0	252	4	US-09-769-482-3
2	1332	98.8	249	4	US-09-769-482-11
3	845.5	62.7	244	4	US-09-769-482-12
4	842.5	62.3	306	4	US-09-091-952A-3
5	791.5	58.7	288	4	US-09-091-952A-4
6	111.5	8.3	1184	4	US-09-266-225D-18
7	111.5	8.3	1185	3	US-09-041-886-23
8	110.5	8.2	732	1	US-08-317-522A-5
9	110.5	8.2	778	1	US-08-439-818A-5
10	110.5	8.2	778	2	US-08-751-965-5
11	110.5	8.2	778	2	US-08-738-975-5
12	110.5	8.2	778	2	US-08-728-626-5
13	110.5	8.2	778	3	US-08-808-599A-5
14	109.5	8.1	1088	4	US-09-233-857-13
15	109.5	8.1	1099	4	US-09-442-100-2
16	109.5	8.1	1099	4	US-08-939-106-2
17	106	7.9	478	4	US-09-252-991A-31455
18	99	7.3	288	4	US-09-252-991A-24917
19	97.5	7.2	902	1	US-08-396-479B-6
20	97.5	7.2	902	1	US-08-818-823-6
21	96.5	7.2	476	3	US-09-189-035-1
22	96.5	7.2	476	3	US-09-382-086-1
23	96.5	7.2	695	2	US-08-701-240-4
24	96.5	7.2	695	3	US-09-138-236-4
25	95.5	7.1	859	1	US-08-395-580-2
26	95.5	7.1	859	5	PCT-US95-02792-2
27	95.5	7.1	1243	2	US-08-557-139-2

28	94.5	7.0	913	1	US-08-445-640-4	Sequence 4, Appli
29	94.5	7.0	913	3	US-08-170-558-4	Sequence 4, Appli
30	94.5	7.0	913	3	US-08-447-314-4	Sequence 4, Appli
31	94.5	7.0	913	3	US-08-445-461-4	Sequence 4, Appli
32	94.5	7.0	919	1	US-08-336-343A-2	Sequence 2, Appli
33	94.5	7.0	1242	4	US-09-508-691-1	Sequence 1, Appli
34	93.5	6.9	534	4	US-09-087-134-11	Sequence 11, Appl
35	93.5	6.9	671	3	US-09-121-321-16	Sequence 16, Appl
36	93.5	6.9	671	3	US-08-933-803A-16	Sequence 16, Appl
37	93.5	6.9	1290	3	US-09-150-460B-6	Sequence 6, Appli
38	93	6.9	776	4	US-09-252-991A-28446	Sequence 28446, A
39	93	6.9	1495	4	US-08-522-726B-1	Sequence 1, Appli
40	93	6.9	1495	4	US-09-337-384-1	Sequence 1, Appli
41	92.5	6.9	522	4	US-09-252-991A-28065	Sequence 28065, A
42	92.5	6.9	556	4	US-09-252-991A-17793	Sequence 17793, A
43	92	6.8	518	3	US-09-113-309-19	Sequence 19, Appl
44	92	6.8	518	3	US-09-521-109-19	Sequence 19, Appl
45	92	6.8	518	4	US-09-562-332-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-769-482-3
; Sequence 3, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-482-3

Query Match	100.0%	Score 1348;	DB 4;	Length 252;
Best Local Similarity	100.0%;	Pred. No. 1.4e-118;		
Matches	252;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRRRDALSSGCLWPS	60	
Db	1	MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQGRRRDALSSGCLWPS	60	
QY	61	ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFQPTYPYLQHEIDLPTISLSDG	120	
Db	61	ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFQPTYPYLQHEIDLPTISLSDG	120	
QY	121	EEPPPYQGCTQLRDPQQLNRESVRAPNRTIFDSDLMDSARLGGPCPPSSNSGIS	180	
Db	121	EEPPPYQGCTQLRDPQQLNRESVRAPNRTIFDSDLMDSARLGGPCPPSSNSGIS	180	
QY	181	ATCYGSGRMGGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW	240	
Db	181	ATCYGSGRMGGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW	240	
QY	241	SKEKDQKQGHPL 252		
Db	241	SKEKDQKQGHPL 252		

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

```

; LEUENIN: 208 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: 1...288
;   OTHER INFORMATION: Clone 22 isoform 2
;                       alternatively spliced
;                       protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4 :
US-09-091-952A-4

```

	Query Match	58.7%;	Score 791.5;	DB 4;	Length 288;	
	Best Local Similarity	64.0%;	Pred. No. 2.3e-66;			
	Matches 160; Conservative	21;	Mismatches 44;	Indels 25;	Gaps 4;	
<hr/>						
QY	2	ALEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSSQRRREDALSSSEGCLWPSE	61	:		
Db	61	SELEFAQIIIVVVTVMVVIVCLLNHYKVSTRSFINRPNQSRREDGL-----	110	:		
<hr/>						
QY	62	STVSGNGIPEPQVYAPRPRTDLAVPPFAQRERFHREQPTYPYLQHEIDLPTTISLSDGE	121	:		
Db	111	-----PQIMHAPRSRDRTAFPSFIQDRFSRFQPTYPVYQHEIDLPTTISLSDGE	160	:		
<hr/>						
QY	122	EPPYQGCTQLQRLDPEQQLELNRESVRAPPNRTIFDSLDMSARL-GGPCPPSSNSGIS	180	:		
Db	161	EPPYQGCTQLQRLDPEQQMELNRESVRAPPNRTIFDSLIDIAMYGGPCPPSSNSGIS	220	:		
<hr/>						
QY	181	ATCYSGGRMEGPPPTYSEVIGHYPGSSFHQQSSGPSSLLEGTRLHHTHIAPLESAAIW	240	:		
Db	221	ASTCSSNGRMEGPPPTYSEVMGHHPGASFLHHQRS---NAHRGSRLQFOQ-NNAESTIVP	276	:		
<hr/>						
QY	241	SKEKDQKQGH	250	:		
Db	277	IKGKDKRKPGN	286	:		

[illegible]

```

RESULT 6
US-09-266-225D-18
; Sequence 18, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266, 225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1184

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-18

Query Match      8.3%; Score 111.5; DB 4; Length 1184;
Best Local Similarity 24.7%; Pred. No. 0.076;
Matches 48; Conservative 14; Mismatches 67; Indels 65; Gaps 7;

QY 52 SSEGCLWPSESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDL 111
Db 379 SSSAAASSSSSSSASFPASQA-----LPSYPH-----SF 411

QY 112 PPTISLDGEEPPPYQGPCTQLRDPEQQLNRESV--RAPPNRTIFDSDLMDSARLGG 169
Db 412 PPPTSLSVSNQPPKYTQP-----SLPSQAVWSQGPPPPPYGRLLANSNAHPG 459

QY 170 PCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFQHQ-----SSGP----- 217
Db 460 PFPPS-----TGAQSTAHPVPVSTHHHHQOQQOQQOQQOQQHHGNSGPPPPPGA 508

QY 218 -PSLLEGTRLHHTH 230
Db 509 FPHPLEGGSSHHAH 522

RESULT 7
US-09-041-886-23
; Sequence 23, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Poly peptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1185 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-041-886-23

Query Match      8.3%; Score 111.5; DB 3; Length 1185;
Best Local Similarity 24.7%; Pred. No. 0.076;
Matches 48; Conservative 14; Mismatches 67; Indels 65; Gaps 7;

52 SSEGCLWPSESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDL 111
380 SSSAAASSSSSSSASFPASQA-----LPSYPH-----SF 412

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-18

Query Match      8.2%; Score 110.5; DB 1; Length 732;
Best Local Similarity 24.5%; Pred. No. 0.05;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;

QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQRERFHRFQPTY--PYLQHEIDLPTI 115
Db 498 SGLPKPCLPEECGEPCPPAEPG-----PP----EAFCSRSEPIPEPSLQEQLEVP 548

QY 116 SLSDGEEPPPYQGPCTQLRDPE----QQLELNRESVRAPP-----NRTIFDSDL 162
Db 549 PPA---EPRPLESCCRSEPIPESSRQEQLEVPPECPAEPRLPESYCRIEPIESSRQ 605

QY 163 DSARLGGPCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFQHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPGPIQPSTQSGS--PPGCPRVE--LGASEPCTLEHRSLESSLPCC 662

QY 221 LEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 663 -----SQWAPATTSLIFFSQ-----HPL 680
```

RESULT 9
US-08-439-818A-5
; Sequence 5, Application US/08439818A
; Patent No. 5654145
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,818A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,522
; FILING DATE: 04-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1563
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 778 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-439-818A-5

Query Match 8.2%; Score 110.5; DB 1; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.054;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;

Qy 67 NGIP-----EPQVYAPRRPTDRLAVPPFAQRRFHRFQPTY--PYLQHEIDLPTI 115
Db 498 SGLPKCLPEECGEPQCPPEPG-----PP-----EAFCSRSEIPEPSLQEQLEVPEPY 548

Qy 116 SLSDGEEPPPYQGPTLQLRDPE-----QQLELNRESVRAPP-----NRTIFDSDL 162
Db 549 PPA---EPRPLESCCRSEIPESSRQEQLEVPEPCPPAEPRPLESYCRIEPEISSRQ 605

Qy 163 DSARLGGPCPPSSNSGISATCYGSGRMGPPPTYSEVIGHYPGSSFQHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPPGLQSTQGSQ-PPGPCRVE--LGASEPCTLEHRSLESSLPCC 662

Qy 221 LEGTRLHHTHIAPLESAAIWSKEDKQKGHPL 252
Db 663 -----SQWAPATTSLIFFSQ-----HPL 680

RESULT 10
US-08-751-965-5
; Sequence 5, Application US/08751965
; Patent No. 5858360
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting

; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,965
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/439,818
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 2252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 778 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-751-965-5

Query Match 8.2%; Score 110.5; DB 2; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.054;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;

Qy 67 NGIP-----EPQVYAPRRPTDRLAVPPFAQRRFHRFQPTY--PYLQHEIDLPTI 115
Db 498 SGLPKCLPEECGEPQCPPEPG-----PP-----EAFCSRSEIPEPSLQEQLEVPEPY 548

Qy 116 SLSDGEEPPPYQGPTLQLRDPE-----QQLELNRESVRAPP-----NRTIFDSDL 162
Db 549 PPA---EPRPLESCCRSEIPESSRQEQLEVPEPCPPAEPRPLESYCRIEPEISSRQ 605

Qy 163 DSARLGGPCPPSSNSGISATCYGSGRMGPPPTYSEVIGHYPGSSFQHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPPGLQSTQGSQ-PPGPCRVE--LGASEPCTLEHRSLESSLPCC 662

Qy 221 LEGTRLHHTHIAPLESAAIWSKEDKQKGHPL 252
Db 663 -----SQWAPATTSLIFFSQ-----HPL 680

RESULT 11
US-08-738-975-5
; Sequence 5, Application US/08738975
; Patent No. 5880267
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/738,975
FILING DATE: herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/439,818
FILING DATE: 05-Dec-1995
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 2251
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 778 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-738-975-5

Query Match 8.2%; Score 110.5; DB 2; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.054;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;
QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQRRFHRFOPTY--PYLQHEIDLPTI 115
Db 498 SGLPKCLPEECGEPCPPAEPG-----PP-----EAFCSRSEPIPEPSLQEQLEVPEPY 548
QY 116 SLSDGEEPPPYQGCTQLQRLDPE-----QQLELNRESVRAPP-----NRTIFDSDLM 162
Db 549 PPA---EPRPLESCCRSEPIPESSRQEQLEVPEPCPPAEPRLPESYCRIEPEIPSSRQ 605
QY 163 DSARLGCGCPPSSNSGISATCYGSGRMEGPPPTTYSEVIGHYPGSSFOHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPGLQPSTQSGG--PPGPCRVE--LGASEPCTLEHRSLESSLPCC 662
QY 221 LEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 663 -----SQWAPATTSLIFSSQ-----HPL 680

RESULT 12
US-08-728-626-5
Sequence 5, Application US/08728626
Patent No. 5910451
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Trophinin and Trophinin-Assisting
TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,626
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/439,818
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: US 08/317,522
FILING DATE: 04-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 1563
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 778 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-728-626-5

Query Match 8.2%; Score 110.5; DB 2; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.054;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;
QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQRRFHRFOPTY--PYLQHEIDLPTI 115
Db 498 SGLPKCLPEECGEPCPPAEPG-----PP-----EAFCSRSEPIPEPSLQEQLEVPEPY 548
QY 116 SLSDGEEPPPYQGCTQLQRLDPE-----QQLELNRESVRAPP-----NRTIFDSDLM 162
Db 549 PPA---EPRPLESCCRSEPIPESSRQEQLEVPEPCPPAEPRLPESYCRIEPEIPSSRQ 605
QY 163 DSARLGCGCPPSSNSGISATCYGSGRMEGPPPTTYSEVIGHYPGSSFOHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPGLQPSTQSGG--PPGPCRVE--LGASEPCTLEHRSLESSLPCC 662
QY 221 LEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 663 -----SQWAPATTSLIFSSQ-----HPL 680

RESULT 13
US-08-808-599A-5
Sequence 5, Application US/08808599A
Patent No. 6111089
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Trophinin, Trophinin-Assisting
TITLE OF INVENTION: Proteins and Methods to Inhibit Implantation
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,599A
FILING DATE: 28-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/317,522
FILING DATE: 04-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/439,818
FILING DATE: 12-MAY-1995
ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 2256
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 778 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-599A-5

Query Match 8.2%; Score 110.5; DB 3; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.054;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;
Qy 67 NGIP-----EPQVYAPRPTDRLAVPPFAQRERFHRFOPTY--PYLQHEIDLPTI 115
Db 498 SGLPKPCLPEEGEPQPCPAEPG-----PP-----EAFCSRSEPIPEPSLQEQLEVPEPY 548
Qy 116 SLSDGEEPPPYQGPCTQLRDPE----QQLELNRESVRAPP-----NRTIFDSDL 162
Db 549 PPA---EPRPLESCCRSEPIPESSRQEQLEVPEPCPPAERPPLSYCRYEPIESSRQ 605
Qy 163 DSARLGGPCPPSSNSGISATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQ--QSSGPPSL 220
Db 606 EQLEVPEPCPPAEPGLOPSTQGSQ--PPGPCRVE--LGASEPCTLEHRSLESSLPBCC 662
Qy 221 LEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 663 -----SQWAPATTSLIFSSQ-----HPL 680

RESULT 14
US-09-233-857-13
Sequence 13, Application US/092333857.
Patent No. 6495353
GENERAL INFORMATION:
APPLICANT: Plowman, Gregory
APPLICANT: Flanagan, Peter
TITLE OF INVENTION: HUMAN ORTHOLOGUES OF WART
FILE REFERENCE: 239/251
CURRENT APPLICATION NUMBER: US/09/233,857
CURRENT FILING DATE: 1999-01-20
EARLIER APPLICATION NUMBER: USSN 60/072,023
EARLIER FILING DATE: 1998-01-21
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13
LENGTH: 1088
TYPE: PRT
ORGANISM: HUMAN
US-09-233-857-13

Query Match 8.1%; Score 109.5; DB 4; Length 1088;
Best Local Similarity 26.6%; Pred. No. 0.11;
Matches 62; Conservative 18; Mismatches 80; Indels 73; Gaps 14;
Qy 28 SHYKLSARSFISRHSGRRREDALSSEGCLWPSESTV---SGNG--IPEPQVY---APPR 79
Db 140 SHYLRCSPALDS--GAGSSRSDSPSHSHHTHQPSRTVGNPGNGGFSPPSGFSEVAPPA 197
Qy 80 PTDR-----LAVPPFAQRERFHRFOPTYPYLQHEI-----DLPTTISLSDGEEPPPP 125
Db 198 PPPRNPTASSAATPP-----PPVPPTSQAYVKRRSPALNNRPPAIA-----PPT 241
Qy 126 YQG--PCTLQ--LRDPEQOQLELNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 181
Db 242 QRGNSPVITQGLKNPQQQLTQQLKSLNLYPGG-----GSGAVVEPPPPYLIQG--- 290
Qy 182 TCYSGGRMEGPP-----PTYSEVIGH--YPGSSFQHQSSGPPS 219

Db 291 ---GAGGAAPPPPPSYTASMQSRQSPQTQSQSDYRKSPSSGIYSATSAGSPS 340
RESULT 15
US-09-442-100-2
Sequence 2, Application US/09442100
Patent No. 6359193
GENERAL INFORMATION:
APPLICANT: Xu, Tian
APPLICANT: Tao, Wufan
APPLICANT: Wang, Weiyl
APPLICANT: Zhang, Sheng
APPLICANT: Yu, Wan
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
TITLE OF INVENTION: GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1099 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-442-100-2

Query Match 8.1%; Score 109.5; DB 4; Length 1099;
Best Local Similarity 26.6%; Pred. No. 0.11;
Matches 62; Conservative 18; Mismatches 80; Indels 73; Gaps 14;
Qy 28 SHYKLSARSFISRHSGRRREDALSSEGCLWPSESTV---SGNG--IPEPQVY---APPR 79
Db 140 SHYLRCSPALDS--GAGSSRSDSPSHSHHTHQPSRTVGNPGNGGFSPPSGFSEVAPPA 197
Qy 80 PTDR-----LAVPPFAQRERFHRFOPTYPYLQHEI-----DLPTTISLSDGEEPPPP 125
Db 198 PPPRNPTASSAATPP-----PPVPPTSQAYVKRRSPALNNRPPAIA-----PPT 241
Qy 126 YQG--PCTLQ--LRDPEQOQLELNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 181
Db 242 QRGNSPVITQGLKNPQQQLTQQLKSLNLYPGG-----GSGAVVEPPPPYLIQG--- 290
Qy 182 TCYSGGRMEGPP-----PTYSEVIGH--YPGSSFQHQSSGPPS 219
Db 291 ---GAGGAAPPPPPSYTASMQSRQSPQTQSQSDYRKSPSSGIYSATSAGSPS 340

Search completed: December 4, 2003, 19:52:39
Job time : 22 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:44:54 ; Search time 18 Seconds
(without alignments)
658.374 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 1348
Sequence: 1 MAELEFVQIIIVVMVMV.....PLESAATWSKEKDQKGHPL 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1340	99.4	287	1	TMEP_HUMAN
2	1172.5	87.0	260	1	TMEP_MOUSE
3	842.5	62.5	306	1	C181_HUMAN
4	114	8.5	1509	1	GSR1_HUMAN
5	113.5	8.4	395	1	UMF1_ARATH
6	113	8.4	628	1	V70K_TYMVC
7	111.5	8.3	1185	1	DRPL_HUMAN
8	110.5	8.2	778	1	TAST_HUMAN
9	104	7.7	830	1	SREC_HUMAN
10	103	7.6	628	1	V70K_TYMV
11	101	7.5	384	1	VASP_CANFA
12	100.5	7.5	1183	1	DRPL_RAT
13	100	7.4	628	1	V70K_TYMVA
14	100	7.4	2440	1	NCR1_HUMAN
15	99.5	7.4	992	1	EBN6_EBV
16	98	7.3	1259	1	AUT2_HUMAN
17	98	7.3	2161	1	SHK1_HUMAN
18	97.5	7.2	902	1	NFC4_HUMAN
19	96	7.1	620	1	EXTN_TOBAC
20	96	7.1	1729	1	TABP_HUMAN
21	96	7.1	1844	1	POLR_TYMVA
22	95.5	7.1	707	1	SFPQ_HUMAN
23	95.5	7.1	859	1	M3KC_HUMAN
24	95	7.0	896	1	CYRB_MOUSE
25	94.5	7.0	276	1	PRPL_HUMAN
26	94.5	7.0	466	1	YL37_CAEL
27	94.5	7.0	913	1	DDR1_HUMAN
28	94.5	7.0	1242	1	TRSL_HUMAN
29	94.5	7.0	1790	1	SEPA_EMENI
30	93.5	6.9	671	1	Z282_HUMAN
31	93.5	6.9	910	1	DDR1_RAT
32	93.5	6.9	2083	1	DYSF_MOUSE
33	93	6.9	244	1	YPHF_BACSU

RESULT 1				
TMEP_HUMAN				
ID	TMEP_HUMAN	STANDARD;	PRT;	287 AA.
AC	Q969W9; Q96B72; Q9UJD3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Transmembrane prostate androgen-induced protein (Solid tumor-associated 1 protein).			
GN	TMEPAI OR PMPAI OR STAG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RX	MEDLINE=20334621; PubMed=10873380;			
RA	Xu L.L., Shanmugam N., Segawa T., Sesterhenn I.A., McLeod D.G.,			
RA	Moul J.W., Srivastava S.;			
RT	"A novel androgen-regulated gene, PMPAI, located on chromosome 20q13			
RT	exhibits high level expression in prostate.";			
RL	Genomics 66:257-263 (2000).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=21453682; PubMed=11568975;			
RA	Rae F.K., Hooper J.D., Nicol D.L., Clements J.A.;			
RT	"Characterization of a novel gene, STAG1/PMPAI, upregulated in renal			
RT	cell carcinoma and other solid tumors.";			
RL	Mol. Carcinog. 32:44-53 (2001).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,			
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,			
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,			
RA	Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,			
RA	Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,			
RA	Lehvaeslaiho M.H., Liversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,			
RA	Marsh V.L., Martin S.L., McConnachie L.J., McIlroy K., McMurray A.A.,			
RA	Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,			
RA	Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,			
RA	Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,			
RA	Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,			
RA	Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,			
RA	Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,			
RA	Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,			
RA	Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,			
RA	Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,			

Not prior.

RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [4]
RP SEQUENCE OF 8-287 FROM N.A. (ISOFORM 1).
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: Interacts with the WW domains of NEDD4 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q969W9-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q969W9-2; Sequence=VSP_006438;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highest expression in prostate. Also expressed
CC in ovary.
CC -!- INDUCTION: By androgen.
CC -!- SIMILARITY: BELONGS TO THE TMEPAI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF224278; AAF86322.1; -;
DR EMBL; AF305616; AAL16781.1; -;
DR EMBL; AF305426; AAL09357.1; -;
DR EMBL; AL035541; CAB55862.1; -;
DR EMBL; BC015918; AAL15918.1; ALT_INIT.
DR Genew; HGNC:14107; TMEPAI.
DR MIM; 606564; -;
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0030521; P:androgen receptor signaling pathway; NAS.
KW Transmembrane; Alternative splicing.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 63 POTENTIAL.
FT DOMAIN 64 287 CYTOPLASMIC (POTENTIAL).
FT SITE 158 161 WW-BINDING (POTENTIAL).
FT SITE 229 232 WW-BINDING (POTENTIAL).
FT VARSPLIC 1 37 MHRMGVNSTAAAGQPNVSTCNCKRSLFQSMEIT ->
FT MA (in isoform 2).
FT /FTId=VSP_006438.
SQ SEQUENCE 287 AA; 31609 MW; 6103473561AE08DA CRC64;

Query Match 99.4%; Score 1340; DB 1; Length 287;
Best Local Similarity 99.2%; Pred. No. 9.7e-94;
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db : |||||
36 ITELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 95
QY 61 ESTVSGNGIPEPOVYAPRPTDRLAVPPEAQRERFHRFQTPYLYQHEIDLPTISLSDG 120
Db |||||
96 ESTVSGNGIPEPOVYAPRPTDRLAVPPEAQRERFHRFQTPYLYQHEIDLPTISLSDG 155
QY 121 EEPPIYQGPCTQLRDPPEQLELNRESVRAPPNRTIFSDMLDSARLGGPCPPSSNSGIS 180
Db |||||
156 EEPPIYQGPCTQLRDPPEQLELNRESVRAPPNRTIFSDMLDSARLGGPCPPSSNSGIS 215
QY 181 ATCYGSGGRMEGPPPTVSEVIGHYPGSSFHQQSSGPPSLLEGTRLHHTHIAPLESAIW 240
Db |||||
216 ATCYGSGGRMEGPPPTVSEVIGHYPGSSFHQQSSGPPSLLEGTRLHHTHIAPLESAIW 275
QY 241 SKEKDKQKGHPL 252
Db |||||
276 SKEKDKQKGHPL 287
RESULT 2
TMEP_MOUSE STANDARD; PRT; 260 AA.
AC Q9D7R2; Q9EQH9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transmembrane prostate androgen-induced protein (Nedd4 WW domain-
DE binding protein 4).
GN TMEPAI OR N4WBP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE OF 59-260 FROM N.A.
RX MEDLINE=20498735; PubMed=11042109;
RA Jolliffe C.N., Harvey K.F., Haines B.P., Parasivam G., Kumar S.;
RT "Identification of multiple proteins expressed in murine embryos as
RT binding partners for the WW domains of the ubiquitin-protein ligase
RT Nedd4.";
RL Biochem. J. 351:557-565(2000).
CC -!- SUBUNIT: Interacts with the WW domains of NEDD4.
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE TMEPAI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AK008976; BAB26001.1; -;
EMBL; AF220208; AAG44247.1; -;
MGD; MGI:1929600; N4wbp4.
GO; GO:0016021; C: integral to membrane; ISS.
GO; GO:0005515; P: protein binding activity; IPI.
GO; GO:0030521; P: androgen receptor signaling pathway; ISS.
Transmembrane.
DOMAIN 1 20 EXTRACELLULAR (POTENTIAL).
TRANSMEM 21 43 POTENTIAL.
DOMAIN 44 260 CYTOPLASMIC (POTENTIAL).
SITE 136 139 WW-BINDING (POTENTIAL).
SITE 207 210 WW-BINDING (POTENTIAL).
SEQUENCE 260 AA; 28715 MW; 31AD07BD16B0D77D CRC64;

Query Match 87.0%; Score 1172.5; DB 1; Length 260;
Best Local Similarity 88.1%; Pred. No. 3.2e-81;
Matches 222; Conservative 11; Mismatches 10; Indels 9; Gaps 3;
1 MALEFVQIIIVVMMVVTCLLSHYKLSARSFISRHSGRRRREDALSGCLWPS 60
18 ITELEFVQIVVIVVMMVVTCLLSHYKLSARSFISRHSGRRRREDALSGCLWPS 77
61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQTPYQLQHEIDLPTISLSDG 120
78 ESTVSG-GMPEPQVYAPRPTDRLAVPPFIQRS---RFQTPYQLQHEIALPTISLSDG 133
121 EEPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180
134 EEPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 193
181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFHQSSGPPSLLEGTRLHHTHIAPLESAI 240
194 ATCYSGGRMEGPPPTYSEVIGHYPGSSFHQSSGPPSLLEGTRLHHTHIAPLESAI 248
241 SKEKDKQKHPL 252
249 NKEKEKQKHPL 260

RESULT 3
C181_HUMAN STANDARD; PRT; 306 AA.
ID C181_HUMAN O15165; O15166; O15167; O15168;
AC O15165; O15166; O15167; O15168;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE protein C18orf1.
GN C18orf1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE=FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=98140124; PubMed=9479497;
RA Yoshikawa T., Sanders A.R., Esterling L.E., Detera-Wadleigh S.D.;
RT "Multiple transcriptional variants and RNA editing in C18orf1, a novel
RL gene with LDLRA and transmembrane domains on 18p11.2";
RL Genomics 47:246-257(1998).
CC -!- SUBCELLULAR LOCATION: Type Ib membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=Alpha-1;
CC IsoId=O15165-1; Sequence=Displayed;
CC Name=Alpha-2;
CC IsoId=O15165-2; Sequence=VSP_006440;

2000

Name=Beta-1;
IsoId=O15165-3; Sequence=VSP_006439;
Name=Beta-2;
IsoId=O15165-4; Sequence=VSP_006439, VSP_006440;
-!- SIMILARITY: BELONGS TO THE TMEPAI FAMILY.
-!- SIMILARITY: Contains 1 LDL-receptor class A domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF009424; AAC52023.1; -;
EMBL; AF009425; AAC52024.1; -;
EMBL; AF009426; AAC52025.1; -;
EMBL; AF009427; AAC52026.1; -;
HSP; P01130; 1AJJ.
Genew; HGNC:1224; C18orf1.
MIM; 606571; -;
GO; GO:0016021; C: integral to membrane; NAS.
InterPro; IPR002172; LDL_receptor_A.
Pfam; PF00057; ldl_recept_a; 1.
SMART; SM00192; LDLa; 1.
PROSITE; PS01209; LDLRA_1; FALSE_NEG.
PROSITE; PS50068; LDLRA_2; 1.
Transmembrane; Alternative splicing.
DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
TRANSMEM 65 85 POTENTIAL.
DOMAIN 86 306 CYTOPLASMIC (POTENTIAL).
DOMAIN 16 48 LDL-RECEPTOR CLASS A.
VARSPPLIC 1 61 MPEAGFQATNAFTCKFTCTSGKCLYGLSLVCNQNDGDN
SDEENCLLVTEHPPPGIFNS -> MAA (in isoform
Beta-1 and isoform Beta-2).
FT VARSPLIC 113 130 /FTId=VSP_006439.
FT Missing (in isoform Alpha-2 and isoform
Beta-2).
FT /FTId=VSP_006440.
SQ SEQUENCE 306 AA; 33900 MW; F48EF66E329201BD CRC64;
Query Match 62.5%; Score 842.5; DB 1; Length 306;
Best Local Similarity 67.2%; Pred. No. 2.1e-56;
Matches 168; Conservative 22; Mismatches 53; Indels 7; Gaps 4;
2 AELEFVQIIIVVMMVVTCLLSHYKLSARSFISRHSGRRRREDALSGCLWPS 61
61 SELEFAQIIIVVVTVMVVIVCLLNHYKVSFTRFNPQSRREDGLPQEGCLWPSD 120
62 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQTPYQLQHEIDLPTISLSDG 121
121 SAAPRLGASE--IMHAPRSRDRFTAPSFQDRFRSQTPYQVQHEIDLPTISLSDG 178
122 EPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLMDSARL-GGPCPPSSNSGIS 180
179 EPPYQGPCTQLRDPEQQLNRESVRAPPNRTIFDSDLDIDIAMYSGGPCPPSSNSGIS 238
181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFHQSSGPPSLLEGTRLHHTHIAPLESAI 240
239 ASTCSSNGRMEGPPPTYSEVGMHHPGASFLHHQRS---NAHRGSRLOFQQ-NNAESTI 294
241 SKEKDKQKH 250
295 IKGDKRKPGN 304
RESULT 4
GSRI_HUMAN STANDARD; PRT; 1509 AA.
ID GSRI_HUMAN
AC Q9NZM4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glioma tumor suppressor candidate region gene 1 protein.
GN GLTSCR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=20175430; PubMed=10708517;
RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,
RA Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,
RA Scheithauer B.W., Louis D.N., Jenkins R.B.;
RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor
RT region.";
RL Genomics 64:44-50(2000).
CC -!- TISSUE SPECIFICITY: Expressed at moderate levels in heart, brain,
CC placenta, skeletal muscle, and pancreas, and at lower levels in
CC lung, liver, and kidney.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF182077; AAF62874.1; -.
DR Genew; HGNC:4332; GLTSCR1.
DR MIM; 605690; -.
FT DOMAIN 37 45 POLY-GLY.
FT DOMAIN 884 889 POLY-PRO.
FT DOMAIN 1214 1225 POLY-SER.
FT DOMAIN 1282 1286 POLY-PRO.
FT DOMAIN 1294 1304 POLY-PRO.
SQ SEQUENCE 1509 AA; 152991 MW; 7C5144F443CE6821 CRC64;

Query Match 8.5%; Score 114; DB 1; Length 1509;
Best Local Similarity 26.9%; Pred. No. 0.53;
Matches 49; Conservative 20; Mismatches 63; Indels 50; Gaps 10;
QY 69 IPEQVYAPP-RPTDR---LAVPPFAQRERFH-----RFQTYP---YLQHEIDLPTIS 116
Db 740 LPSHPTRPPSRPPSRPQSVSRPP--SEPLHPCPPPPQAPPTLGFIVIQNLGVPPAS 797
QY 117 LSDGEEPPPYQGCTLQLRDPQLNRESVRAPPNRTIFSDLMDSARLGGPCPPSSN 176
Db 798 NPAPTAPGPPQPLRPQSQPPGGL-----PPAPHL-----PPSST 833
QY 177 SGISATCYSGGRMEGPPPTYSVIGHYPGSSFQHQSSGPPSLLEGRLHHTHIAPLES 236
Db 834 SSAVASSSETSSRL--PAPTPSDQLQFPSPSQGPHKSPTPPTL-----HLVP-EP 881
QY 237 AA 238
Db 882 AA 883

RESULT 5
UMPI_ARATH STANDARD; PRT; 395 AA.
ID UMPI_ARATH
AC Q9LKAS;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Unknown mitochondrial protein At3g15000.
GN AT3G15000 OR K15M2.14.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
RT and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGECC).";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 169-179, AND SUBCELLULAR LOCATION.
RC TISSUE=Leaf, and Stem;
RX MEDLINE=21608403; PubMed=11743114;
RA Kruft V., Eubel H., Jaensch L., Werhahn W., Braun H.-P.;
RT "Proteomic approach to identify novel mitochondrial proteins in
RT Arabidopsis.";
RL Plant Physiol. 127:1694-1710(2001).
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP00370; BAA97063.1; -.
DR EMBL; AF428427; AAL16196.1; -.
KW Mitochondrion.
SQ SEQUENCE 395 AA; 42869 MW; D15E5CCC89E2A33C CRC64;

Query Match 8.4%; Score 113.5; DB 1; Length 395;
Best Local Similarity 23.7%; Pred. No. 0.13;
Matches 46; Conservative 12; Mismatches 61; Indels 75; Gaps 8;
QY 45 RRREDALSSGCLWPSESTVSGNGIPEQV--YAPRPTDRLAVPPFAQRERFHRFOPTY 102
Db 232 RRRENMAAGP---PPQRPMPGPPPPPHHIGGSAPPPHMGGSAPP-----P 274
QY 103 PYLQHEIDLPTISLSDGEEPPPYQGCTLQLRDPQLNRESVRAPPNRTIFSDLM 162
Db 275 PHMGQNYGPPPPNNMGGRRHPPPYGAP-----PQN----- 304
QY 163 DSARLGGPCPPSSNSGISATCYGS-----GGRMEGPPPTYSEVI-----G 202
Db 305 ---NMGGPRPPQNYGGTTPPNYGGAPPANMGGAPPNNYGG---GPPPYGAVPPPYGG 358
QY 203 HYPGSSFQHQSSG 216
Db 359 APPQNNYQQQSG 372

RESULT 6
V70K_TYMVC STANDARD; PRT; 628 AA.
ID V70K_TYMVC
AC P28478;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE 69 kDa protein.
OS Turnip yellow mosaic virus (isolate TYMC).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.
OX NCBI_TaxID=31751;
RN [1]
RP SEQUENCE FROM N.A.

DR Genew; HGNC:3033; DRPLA.
DR MIM; 607462; -.
DR MIM; 125370; -.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0005515; F:protein binding activity; TAS.
DR GO; GO:0007417; P:central nervous system development; TAS.
DR InterPro; IPR002951; Atrophin.
DR Pfam; PF03154; Atrophin-1; 2.
DR PRINTS; PR01222; ATROPHIN.
KW Triplet repeat expansion; Polymorphism.
FT DOMAIN 73 82 SER/GLU-RICH (MIXED CHARGE).
FT DOMAIN 302 305 POLY-PRO.
FT DOMAIN 376 382 POLY-SER.
FT DOMAIN 386 397 POLY-SER.
FT DOMAIN 442 447 POLY-PRO.
FT DOMAIN 479 483 POLY-HIS.
FT DOMAIN 484 497 POLY-GLN.
FT DOMAIN 504 507 POLY-PRO.
FT DOMAIN 564 574 POLY-SER.
FT DOMAIN 704 707 POLY-PRO.
FT DOMAIN 802 815 ARG/ALA-RICH (MIXED CHARGE).
FT DOMAIN 816 827 ARG/GLU-RICH (MIXED CHARGE).
FT DOMAIN 925 934 ARG/GLU-RICH (MIXED CHARGE).
FT CONFLICT 94 94 MISSING (IN REF. 3).
FT CONFLICT 333 333 H -> Y (IN REF. 1).
FT CONFLICT 339 339 M -> I (IN REF. 3).
FT CONFLICT 541 541 P -> T (IN REF. 6).
FT CONFLICT 1028 1028 A -> G (IN REF. 1).
SQ SEQUENCE 1185 AA; 124773 MW; 5135744CEE491C18 CRC64;

Query Match 8.3%; Score 111.5; DB 1; Length 1185;
Best Local Similarity 24.7%; Pred. No. 0.62;
Matches 48; Conservative 14; Mismatches 67; Indels 65; Gaps 7;
QY 52 SSEGLWPSESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDL 111
Db 380 SSSAAASSSSSSSSASPPASQA-----LPSYPH-----SF 412
QY 112 PPTISLDGEEPPPYQGCTQLQRDPEQQLNRESV--RAPNRTIFDSDLMSARLGG 169
Db 413 PPPTSLSVSNQPKYTQP-----SLPSQAVWSQGGPPPPPYGRLLANSNAHPG 460
QY 170 PCPPSSNSGISATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQ-----SSGP----- 217
Db 461 PFPPPS-----TGAQSTAHPPVSTHHHHHQQQQQQQQQQQQHGNSSGPPPPGA 509
QY 218 -PSLLEGTRLHHTH 230
Db 510 FPHPLEGGSSHHAH 523

RESULT 8
TAST_HUMAN STANDARD; PRT; 778 AA.
AC Q12815;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Trophinin-associated protein (tastin) (Trophinin-assisting protein).
GN TROAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95278733; PubMed=7758945;
RA Fukuda M.N., Sato T., Nakayama J., Klier G., Mikami M., Aoki D.,
RA Nozawa S.;
RT "Trophinin and tastin, a novel cell adhesion molecule complex with
RT potential involvement in embryo implantation."
RL Genes Dev. 9:1199-1210(1995).

CC -!- FUNCTION: COULD BE INVOLVED WITH BYSTIN AND TROPHININ IN A CELL
CC ADHESION MOLECULE COMPLEX THAT MEDIATES AN INITIAL ATTACHMENT OF
CC THE BLASTOCYST TO UTERINE EPITHELIAL CELLS AT THE TIME OF THE
CC EMBRYO IMPLANTATION.
CC -!- SUBUNIT: DIRECTLY BINDS BYSTIN, AND INDIRECTLY TROPHININ.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: STRONG EXPRESSION AT IMPLANTATION SITES. WAS
CC EXCLUSIVELY LOCALIZED TO THE APICAL SIDE OF THE
CC SYNCYTOTROPHOBLAST. ALSO FOUND IN MACROPHAGES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U04810; AAA79333.1; -.
DR PIR; I38487; I38487.
DR Genew; HGNC:12327; TROAP.
DR MIM; 603872; -.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
KW Cell adhesion; Repeat.
FT DOMAIN 504 687 CYS-RICH.
FT DOMAIN 516 647 4 X 33 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 516 548 1.
FT REPEAT 549 581 2.
FT REPEAT 582 614 3.
FT REPEAT 615 647 4.
SQ SEQUENCE 778 AA; 83758 MW; 52BAB17165672AC0 CRC64;

Query Match 8.2%; Score 110.5; DB 1; Length 778;
Best Local Similarity 24.5%; Pred. No. 0.46;
Matches 52; Conservative 28; Mismatches 77; Indels 55; Gaps 12;
QY 67 NGIP-----EPQVYAPRPTDRLAVPPFAQRERFHRFQPTY--PYLQHEIDLPTTI 115
Db 498 SGLPKCLPECEGEPQCPPEPG-----PP-----EAFCRSEPEIPEPSLQEQLEVEPEY 548
QY 116 SLSDGEEPPPYQGCTQLQRDPE---QQLELNRESVRAPP-----NRTIFDSDL 162
Db 549 PPA---EPRPLESCCRSEPEIPESSRQEQLEVEPECPPEPRPLESYCRIEPEIPESSRQ 605
QY 163 DSARLGGPCPPSSNSGISATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQ--QSSGPPSL 220
Db 606 EQLEVEPECPPEPGPLQPSQTQSG--PPGPCRVE--LGASEPCTLEHRSLESSLPPCC 662
QY 221 LEGTRLHHTHIAPIESAAIWSKEKDQKHPL 252
Db 663 -----SQWAPATTSLIFSSQ-----HPL 680

RESULT 9
SREC_HUMAN STANDARD; PRT; 830 AA.
AC Q14162; O43701;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Endothelial cells scavenger receptor precursor (Acetyl LDL receptor)
DE (Scavenger receptor class F member 1).
GN SCARF1 OR SREC OR KIAA0149.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical vein endothelial cells;
RX MEDLINE=98058897; PubMed=9395444;
RA Adachi H., Tsujimoto M., Arai H., Inoue K.;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U31777; AAA80337.1; --
EMBL; X89453; CAA61623.1; --
InterPro; IPR002951; Atrophin.
Pfam; PF03154; Atrophin-1; 2.
PRINTS; PRO1222; ATROPHIN.
DOMAIN 165 171 POLY-PRO.
DOMAIN 303 306 POLY-PRO.
DOMAIN 377 383 POLY-SER.
DOMAIN 387 391 POLY-SER.
DOMAIN 440 446 POLY-PRO.
DOMAIN 477 480 POLY-HIS.
DOMAIN 481 489 POLY-GLN.
DOMAIN 502 505 POLY-PRO.
DOMAIN 562 572 POLY-SER.
DOMAIN 702 705 POLY-PRO.
CONFLICT 455 455 N -> S (IN REF. 2).
CONFLICT 594 594 F -> L (IN REF. 2).
CONFLICT 689 689 P -> R (IN REF. 2).
CONFLICT 717 717 T -> M (IN REF. 2).
CONFLICT 737 737 A -> V (IN REF. 2).
CONFLICT 965 965 MISSING (IN REF. 2).
SEQUENCE 1183 AA; 124778 MW; 7FB9928DCADF9B1F CRC64;

Query Match 7.5%; Score 100.5; DB 1; Length 1183;
Best Local Similarity 22.6%; Pred. No. 4.1;
Matches 35; Conservative 14; Mismatches 45; Indels 61; Gaps 7;
QY 100 PTYPYLQHEIDLPTTISLSDGHEPPPYQGPCTLQLRDPQQLNRESVRAPPNRTIFDS 159
Db 404 PSYPH-----SFPPPTSMVSNGPKYTP-----SLPSQAVWSQ 438
QY 160 DLMSARLGGPCPPSSNSGISATCYGSGGRMEGP-PPTYSEVIGHYPGSSFQHQ----- 213
Db 439 -----GPPPPPPPYG---RLLPNNNTHPGFPPTGGQSTAHPPAPAHHHHQQQQQP 486
QY 214 -----SSGP-----PSLLEGTRLHHTH 230
Db 487 QPQPQFQHHHNGSGPPPPGAYPHPLESSNSHHAH 521

RESULT 13
V70K_TYMVA
ID V70K_TYMVA STANDARD; PRT; 628 AA.
AC P20131;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE 69 kDa protein.
OS Turnip yellow mosaic virus (Australian isolate).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymovirus.
OX NCBI_TaxID=12155;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90021184; PubMed=2800335;
RA Keese P., Mackenzie A., Gibbs A.;
RT "Nucleotide sequence of the genome of an Australian isolate of turnip yellow mosaic tymovirus.";
RL Virology 172:536-546(1989).
CC -!- FUNCTION: NOT KNOWN.
CC -!- SIMILARITY: TO 65 TO 70 kDa PROTEIN FROM OTHER TYMOVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; J04373; AAA46591.1; --
DR PIR; JQ0110; JQ0110.
DR InterPro; IPR004935; Tymo 45 70kDa.
DR Pfam; PF03251; Tymo_45kd_70kd; 1.
SQ SEQUENCE 628 AA; -68740 MW; 67CD342A09161D96 CRC64;

Query Match 7.4%; Score 100; DB 1; Length 628;
Best Local Similarity 25.0%; Pred. No. 2.2;
Matches 55; Conservative 19; Mismatches 92; Indels 54; Gaps 10;

QY 40 RSHQRRREDALSSGCLWPSESTVSGNG-----IPEQVYAPRPTDRLAVPPF 89
Db 394 RH---RRSHPLLPPPAALPPIAYTSGRGKIHSLPKGALPKGPPPPRRRLPSPATPPQ 450
QY 90 AQRRFRHFQPTYPYLQHEIDLPTTISLSDGEE---PPPYQGPCTLQLRDPQQLNLR 145
Db 451 SPLRDLGR-TPSFP-----TPPKTSTRAPESCIAAPPTDIAPL-----DSDPVLVS-R 496
QY 146 ESVRAPPNRTIFDSLMDLSARLGGPCPPSSNSGISATCYGSGGRMEGPPPTYSEVIGHYP 205
Db 497 TEVHAPERRTFMDPEALRAALAGLSPPLS-VGIIRTASQTVLPTNSPSPT----- 546
QY 206 GSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKD 245
Db 547 ----RHLPTSSPWILQ-----SPVGEDAIVDSEDD 573

RESULT 14
NCRI_HUMAN
ID NCRI_HUMAN STANDARD; PRT; 2440 AA.
AC 075376; Q9UPV5; Q9UQ18;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nuclear receptor co-repressor 1 (N-Cor1) (N-Cor).
GN NCOR1 OR KIAA1047.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98393736; PubMed=9724795;
RA Wang J., Hoshino T., Redner R.L., Kajigaya S., Liu J.M.;
RT "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-Cor/msin3/HDAC1 complex.";
RT Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998).
RN [2]
RP SEQUENCE OF 782-2440 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
RN [3]
RP SEQUENCE OF 974-2440 FROM N.A.
RX MEDLINE=99375328; PubMed=10444336;
RA Nagaya T., Chen K.-S., Fujieda M., Ohmori S., Richer J.K., Horwitz K.B., Lupski J.R., Seo H.;
RT "Localization of the human nuclear receptor co-repressor (hN-Cor) gene between the CMT1A and the SMS critical regions of chromosome 17p11.2.";
RL Genomics 59:339-341(1999).
CC -!- FUNCTION: Mediates the transcriptional repression activity of some nuclear receptors by promoting chromatin condensation, thus preventing access of the basal transcription.
CC -!- SUBUNIT: Interacts with HDAC7 (By similarity). Forms a large

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:48:04 ; Search time 35 seconds
(without alignments)
1857.979 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 1348
Sequence: 1 MALEFVQIIIIIVVMVMV.....PLESAAIWSKEKDQKGHPL 252

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282	95.1	237	4 Q8NER4	Q8ner4 homo sapien
2	820.5	60.9	292	11 Q8BYE2	Q8bye2 mus musculu
3	817.5	60.6	306	11 Q8BWJ4	Q8bwj4 mus musculu
4	662.5	49.1	200	11 Q8R5E2	Q8r5e2 mus musculu
5	252	18.7	53	4 Q8BYL8	Q8byl8 homo sapien
6	244	18.1	84	4 Q9NTR9	Q9ntr9 homo sapien
7	126	9.3	593	4 Q8IZL1	Q8izl1 homo sapien
8	125.5	9.3	494	4 Q96AI8	Q96ai8 homo sapien
9	125.5	9.3	607	4 Q8IZ34	Q8iz34 homo sapien
10	125.5	9.3	634	4 Q8N6W3	Q8n6w3 homo sapien
11	125.5	9.3	634	4 Q75114	Q75114 homo sapien
12	123.5	9.2	496	10 Q8VZR8	Q8vzr8 arabidopsis
13	123.5	9.2	508	10 Q9M0H8	Q9m0h8 arabidopsis
14	123	9.1	462	11 Q8CFS8	Q8cfs8 mus musculu
15	122	9.1	357	10 Q8L7D7	Q8l7d7 arabidopsis
16	122	9.1	530	10 Q9LYK5	Q9lyk5 arabidopsis

17	118	8.8	291	6	Q97928	Q97928 pongo pygma
18	118	8.8	593	4	Q9H801	Q9h801 homo sapien
19	117.5	8.7	286	6	Q97957	Q97957 macaca fasc
20	117.5	8.7	301	6	Q97923	Q97923 hylobates 1
21	115	8.5	299	6	Q97922	Q97922 gorilla gor
22	115	8.5	471	4	Q8N684	Q8n684 homo sapien
23	113.5	8.4	729	5	Q18660	Q18660 drosophila
24	112	8.3	471	11	Q8BTV2	Q8btv2 mus musculu
25	111.5	8.3	1182	4	Q99495	Q99495 homo sapien
26	111	8.2	727	10	Q9XIL9	Q9xil9 arabidopsis
27	110.5	8.2	300	10	Q8LIW2	Q8liw2 oryza sativ
28	110.5	8.2	778	4	Q8N5B2	Q8n5b2 homo sapien
29	109.5	8.1	596	5	Q8MZ04	Q8mz04 drosophila
30	109.5	8.1	610	5	Q8IPK6	Q8ipk6 drosophila
31	109.5	8.1	729	5	Q9VMS4	Q9vms4 drosophila
32	109.5	8.1	1099	5	Q24096	Q24096 drosophila
33	109.5	8.1	1099	5	Q24590	Q24590 drosophila
34	109.5	8.1	1105	5	Q9VA38	Q9va38 drosophila
35	109	8.1	1190	4	Q99621	Q99621 homo sapien
36	108.5	8.0	994	5	Q9BLN0	Q9bln0 theileria a
37	108.5	8.0	1167	5	Q9Y067	Q9y067 theileria a
38	108.5	8.0	3469	5	Q9U4I2	Q9u4i2 drosophila
39	108.5	8.0	3604	5	Q9VYK0	Q9vyk0 drosophila
40	107.5	8.0	1118	5	Q9VR13	Q9vr13 drosophila
41	107	7.9	313	6	Q97927	Q97927 pan paniscu
42	107	7.9	809	5	Q9NAN8	Q9nan8 caenorhabdi
43	107	7.9	2357	5	Q9U1M8	Q9u1m8 dictyosteli
44	106.5	7.9	358	10	Q9MAH5	Q9mah5 arabidopsis
45	106.5	7.9	939	10	Q8LJ66	Q8lj66 oryza sativ

ALIGNMENTS

RESULT 1

Q8NER4 ID Q8NER4 PRELIMINARY; PRT; 237 AA.
AC Q8NER4;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE PMEPAL variant A protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Brunschwig E.B., Wilson K., Mack D., Dawson D., Lawrence E.,
RA Willson J.K.V., Lu S., Nosrati A., Swinler S., Beard L.,
RA Lutterbaugh J.D., Willis J., Platzer P., Markowitz S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY128643; AAM89277.1; --
SQ SEQUENCE 237 AA; 26201 MW; A44A274EAABFD930 CRC64;

Query Match		95.1%	Score 1282;	DB 4;	Length 237;
Best Local Similarity		100.0%	Pred. No. 4.9e-117;		
Matches 237;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	16	MMVMVVITCLLSHYKLSARSFISRH	SOQRRREDALISSEGCLWPSESTVSGNGIPEQVY	75	
Db	1	MMVMVVITCLLSHYKLSARSFISRH	SOQRRREDALISSEGCLWPSESTVSGNGIPEQVY	60	
QY	76	APPRPTDRLAVPPFAQERFHRFQPT	YPYLOHEIDLPTTISLSDGEEPPPYQGPTQLQR	135	
Db	61	APPRPTDRLAVPPFAQERFHRFQPT	YPYLOHEIDLPTTISLSDGEEPPPYQGPTQLQR	120	
QY	136	DPEQQLNRESVRAPNRTIFDSDLMD	SARLGPGPPSSNSGISATCYGSGGRMEGPPP	195	
Db	121	DPEQQLNRESVRAPNRTIFDSDLMD	SARLGPGPPSSNSGISATCYGSGGRMEGPPP	180	
QY	196	TYSEVIGHYPGSSFQHOQSSGPPSL	LEGTRLHHTHIAPLESAAIWSKEKDQKGHPL	252	

Db 181 TYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDKQKGHPL 237

RESULT 2

Q8BYE2

ID Q8BYE2 PRELIMINARY; PRT; 292 AA.

AC Q8BYE2;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Protein C18ORF1 homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Thymus;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK040289; BAC30562.1; --

SQ SEQUENCE 292 AA; 32521 MW; F780D391984D66E3 CRC64;

Query Match 60.9%; Score 820.5; DB 11; Length 292;

Best Local Similarity 63.6%; Pred. No. 7.1e-72;

Matches 164; Conservative 24; Mismatches 45; Indels 25; Gaps 5;

QY 2 AELEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSSEGCLWPSE 61

Db 47 AELEFAQILIIVVVTVVVVVVCLLNHYKVSFTRFNQSQSQEQEDGLQPEGSLWPSD 106

QY 62 STVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQTPYLYQHEIDLPTISLSDGE 121

Db 107 SSVQRPQASE--IMCAPRGDRFTTPSFQIRDPFSRFQTPYLYQHEIDLPTISLSDGE 164

QY 122 EPPYQGPCTQLRDPEQQLELNRESVRAPPNRTIFDSDLMD-SARLGGPCPPSSNSGIS 180

Db 165 EPPYQGPCTQLRDPEQQLELNRESVRAPPNRTVFDSLIDISMYNGGCPPPSSHSGIS 224

QY 181 ATCYGSGGRMEGPPPTYSYEVIGHYPGSSFOHQSSG-----PPSLLEGTRLHHTHI 231

Db 225 AATCSSNGRMEGPPPTYSYEVGMHYPGTSTFFHHQHSNTHRGSRPQFPNNSEGT-----I 278

QY 232 APLESAIWSKEKDKQKG 249

Db 279 VPI-----KGKDRKPG 289

RESULT 3

Q8BWJ4

ID Q8BWJ4 PRELIMINARY; PRT; 306 AA.

AC Q8BWJ4;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Protein C18ORF1 homolog.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Heart;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK052338; BAC34943.1; --

SQ SEQUENCE 306 AA; 33857 MW; 043C2260AEE8729F CRC64;

Query Match 60.6%; Score 817.5; DB 11; Length 306;

Best Local Similarity 63.2%; Pred. No. 1.5e-71;

Matches 163; Conservative 25; Mismatches 45; Indels 25; Gaps 5;

QY 2 AELEFVQIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSSEGCLWPSE 61

Db 61 SELEFAQILIIVVVTVVVVVVCLLNHYKVSFTRFNQSQSQEQEDGLQPEGSLWPSD 120

QY 62 STVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQTPYLYQHEIDLPTISLSDGE 121

Db 121 SSVQRPQASE--IMCAPRGDRFTTPSFQIRDPFSRFQTPYLYQHEIDLPTISLSDGE 178

QY 122 EPPYQGPCTQLRDPEQQLELNRESVRAPPNRTIFDSDLMD-SARLGGPCPPSSNSGIS 180

Db 179 EPPYQGPCTQLRDPEQQLELNRESVRAPPNRTVFDSLIDISMYNGGCPPPSSHSGIS 238

QY 181 ATCYGSGGRMEGPPPTYSYEVIGHYPGSSFOHQSSG-----PPSLLEGTRLHHTHI 231

Db 239 AATCSSNGRMEGPPPTYSYEVGMHYPGTSTFFHHQHSNTHRGSRPQFPNNSEGT-----I 292

QY 232 APLESAIWSKEKDKQKG 249

Db 293 VPI-----KGKDRKPG 303

RESULT 4

Q8R5E2

ID Q8R5E2 PRELIMINARY; PRT; 200 AA.

AC Q8R5E2;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Similar to chromosome 18 open reading frame 1 (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC022716; AAH22716.1; --

FT NON TER 1

SQ SEQUENCE 200 AA; 22117 MW; 14A78F6E8A4C09A5 CRC64;

Query Match 49.1%; Score 662.5; DB 11; Length 200;

Best Local Similarity 62.7%; Pred. No. 1.2e-56;

Matches 133; Conservative 15; Mismatches 39; Indels 25; Gaps 5;

QY 48 EDALSSSEGCLWPSESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRERFHRFQTPYLYQH 107

Db 1 EDGLQPEGSLWPSDSSVQRPQASE--IMCAPRGDRFTTPSFQIRDPFSRFQTPYLYQH 58

QY 108 EIDLPTISLSDGEEPPYQGPCTQLRDPEQQLELNRESVRAPPNRTIFDSDLMD-SAR 166

Db 59 EIDLPTISLSDGEEPPYQGPCTQLRDPEQQLELNRESVRAPPNRTVFDSLIDISMY 118

QY 167 LGGPCPPSSNSGISATCYGSGGRMEGPPPTYSYEVIGHYPGSSFOHQSSG-----P 217

Db 119 NGGCPPPSSHSGISAATCSSNGRMEGPPPTYSYEVGMHYPGTSTFFHHQHSNTHRGSRPQFQ 178

QY 218 PSLLLEGTRLHHTHIAPLESAAIWSKEKDKQKG 249

Db 179 PNNSEGT-----IVPI-----KGKDRKPG 197

RESULT 5

Q9BYL8

ID Q9BYL8 PRELIMINARY; PRT; 53 AA.

AC Q9BYL8;

Qy 133 QLRDEQ-----QLELNRESVRAAPPNRTIFDSDLMDLSARLGGPCPPSSNSG 178
Db 422 PQSPENPGQPSMGIDIASAPALQOVRTSAGSPANQS-----PTSPVSNQG 467
Qy 179 ISATCYGSGRMEGPPPTTYSEVIGHYPGSSFQHQ 213
Db 468 FS-----PGSSPQHTSLGVSFGDAYYEQQ 492

RESULT 12

Q8VZR8 ID Q8VZR8 PRELIMINARY; PRT; 496 AA.
AC Q8VZR8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative proline-rich protein.
GN AT4G28300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.
RT "Full length cDNA of gene At4g28300 (GI:15235249).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY063895; AAL36251.1; -.
DR EMBL; AY096387; AAM20028.1; -.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
SQ SEQUENCE 496 AA; 54106 MW; 1BF54098B9A79429 CRC64;

Query Match 9.2%; Score 123.5; DB 10; Length 496;
Best Local Similarity 26.4%; Pred. No. 0.0015;
Matches 51; Conservative 18; Mismatches 71; Indels 53; Gaps 10;
Qy 70 PEPQVYAPPRPTDRLAVPP-----FAQRERFHRFQPTYP-----YLQHEIDLPTISLS 118
Db 259 PPSQLQAPPAQSQFMPPPPAPSHPSAQTQSFQYQNNWPPQPQARPPQSSGGYPTYSPP 318
Qy 119 DGEEPP-----PYQGPTLQLRDPEQQLNRESVRAPPNRTIFDSDLMDLSAR 166
Db 319 PGNQPPVESLPSSMQMSPYSGP-----PQSQMAYGYGAAPPQAPPQQTQKMSYSPQ 371
Qy 167 LG-----GPCPPSSNSGISCATCYGSGGRMEGPP--PTYSEVIGHY-----PGSSFQHQ 212
Db 372 TGDGYLPSGPPPP--SGYANAMY-EGGRMQYPPPPQPPQOQQQAAHYLQGPQGGGYSPPH 427
Qy 213 QSSG-----PPSL 220
Db 428 QAGGNIGAPPVL 440

RESULT 13

Q9M0H8 ID Q9M0H8 PRELIMINARY; PRT; 508 AA.
AC Q9M0H8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Predicted proline-rich protein.
GN AT4G28300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161572; CAB79632.1; -.
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
SQ SEQUENCE 508 AA; 55574 MW; 0C5C6C09926A1164 CRC64;

Query Match 9.2%; Score 123.5; DB 10; Length 508;
Best Local Similarity 26.4%; Pred. No. 0.0015;
Matches 51; Conservative 18; Mismatches 71; Indels 53; Gaps 10;

Qy 70 PEPQVYAPPRPTDRLAVPP-----FAQRERFHRFQPTYP-----YLQHEIDLPTISLS 118
Db 271 PPSQLQAPPAQSQFMPPPPAPSHPSAQTQSFQYQNNWPPQPQARPPQSSGGYPTYSPP 330
Qy 119 DGEEPP-----PYQGPTLQLRDPEQQLNRESVRAPPNRTIFDSDLMDLSAR 166
Db 331 PGNQPPVESLPSSMQMSPYSGP-----PQSQMAYGYGAAPPQAPPQQTQKMSYSPQ 383
Qy 167 LG-----GPCPPSSNSGISCATCYGSGGRMEGPP--PTYSEVIGHY-----PGSSFQHQ 212
Db 384 TGDGYLPSGPPPP--SGYANAMY-EGGRMQYPPPPQPPQOQQQAAHYLQGPQGGGYSPPH 439
Qy 213 QSSG-----PPSL 220
Db 440 QAGGNIGAPPVL 452

RESULT 14

Q8CFS8 ID Q8CFS8 PRELIMINARY; PRT; 462 AA.
AC Q8CFS8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to hypothetical protein FLJ12529.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038812; AAH38812.1; -.
KW Hypothetical protein.
SQ SEQUENCE 462 AA; 51056 MW; 42D511ECC74E228E CRC64;
Query Match 9.1%; Score 123; DB 11; Length 462;
Best Local Similarity 25.4%; Pred. No. 0.0015;

Matches	50;	Conservative	17;	Mismatches	70;	Indels	60;	Gaps	8;
Qy	27	LSHYKLSARSFISRH	SQRRREDALSS	EGCLWPSESTVSGN-----	67				
Db	165	LSQFEAQARKRIPPR	AHSRDSSD--SAD	GRATPSENLPSSARVDKPPSVLPYFNRP	PSA 222				
Qy	68	---GIP	EPQVYAPRPTDRL	AVPPFAQRERFHR	FQPTYPLQHEIDL	PPTISLSDGEEP	123		
Db	223	LPLMGLPPPIPPPP	LSSSFGVPPPPG	IHYQHLMPPPP-----	RLPPHLAV-----P	271			
Qy	124	PPYQGPCTLQLRD	PEQOQLELNRES	VRAPPNRTIFDS	DLMDSARLGGPCPP	SSNSGISATC	183		
Db	272	PPGAIPPALHL--	-----NP	AFPPPNATV-----	GP-PPDTYMKASTPY	308			
Qy	184	YSGGRMEGPPP	-TYSE 199						
Db	309	NHGRDSDGPP	PSTVSE 325						

RESULT 15									
ID	Q8L7D7	PRELIMINARY;	PRT;	357	AA.				
AC	Q8L7D7;								
DT	01-OCT-2002	(TrEMBLrel. 22, Created)							
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)							
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)							
DE	Hypothetical protein.								
GN	AT5G14540.								
OS	Arabidopsis thaliana (Mouse-ear cress).								
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;								
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsis.								
OX	NCBI_TaxID=3702;								
RN	{1}								
RP	SEQUENCE FROM N.A.								
RA	Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,								
RA	Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,								
RA	Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,								
RA	Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,								
RA	Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,								
RA	Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,								
RA	Theologis A., Davis R.W.;								
RL	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AY136314; AAM96980.1; -								
DR	InterPro; IPR002965; P rich_extensn.								
DR	InterPro; IPR000449; UBA_domain.								
DR	PRINTS; PR01217; PRICHEXTENSN.								
KW	Hypothetical protein.								
SQ	SEQUENCE	357	AA;	38339	MW;	499FDF163DDF8030.CRC64;			
Query Match									
Best Local Similarity 9.1%; Score 122; DB 10; Length 357;									
Matches 48; Conservative 8; Mismatches 71; Indels 38; Gaps 8;									
QY	65	SGNGIPEPQV---	YAPPRPTDRLAVPPFAQRERFHR	FQPTYPLQHEIDL	PPTISLSDGE	121			
Db	111	SGQSQPPPTIQPPYQPPPTQSLHQPPYQPPQ----	QPQYPQ-----QP	PPQLQHP	SGY 161				
QY	122	EP--PPYQGPCTLQLRDPEQOQLELNRES	VRAPPNRTIFDS	DLMDSARLGGPCPPS	-----	174			
Db	162	NPEEPY-----	PQSYPPN--P	RPQPPSHPPGSA	PSQQYNNAPPTPP	SMYDGP 209			
QY	175	---SNSGISATCYGSGGRMEGPPPTTYSEVIGHYPGSS	FQHQSSG	216					
Db	210	GGRSNSGFPSGYSPE	SYPTGPPSQY----	GNTPSVKP	THQSGSG	250			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 19:44:19 ; Search time 41 Seconds
(without alignments)
975.588 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 1348
Sequence: 1 MAPLEFVQIIIVVMVMV.....PLESAAIWSKEKDQKQGHPL 252

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1348	100.0	252	21 AAB18449	A human TANGO 261
2	1348	100.0	252	21 AAB18449	Neuron-associated
3	1348	100.0	252	23 ABG61805	Prostate cancer-as
4	1348	100.0	285	24 ABP75660	Human secretory po
5	1345	99.8	252	21 AAB18461	A human TANGO 261
6	1345	99.8	252	21 AAB18462	A human TANGO 261
7	1345	99.8	252	21 AAB18463	A human TANGO 261
8	1340	99.4	287	23 AAU78231	Human Mechanically
9	1282	95.1	237	22 AAM38712	Human polypeptide

10	1172.5	87.0	274	23	AAU78236	Mouse Mechanically
11	1133	84.1	207	22	AAM40498	Human polypeptide
12	1066.5	79.1	217	21	AAB18450	A murine TANGO 261
13	1065.5	79.0	241	23	ABG77052	Prostate specific
14	1063.5	78.9	217	21	AAB18464	A murine TANGO 261
15	1063.5	78.9	217	21	AAB18465	A murine TANGO 261
16	1063.5	78.9	217	21	AAB18466	A murine TANGO 261
17	844.5	62.6	269	21	AAU73422	Human secreted pro
18	843.5	62.6	306	19	AAW60673	Manic-depressive i
19	842.5	62.5	334	22	ABG13117	Novel human diagno
20	796	59.1	261	22	ABG20316	Novel human diagno
21	792.5	58.8	288	19	AAW60674	Manic-depressive i
22	791.5	58.7	316	22	ABG13118	Novel human diagno
23	765.5	56.8	235	22	ABG13123	Novel human diagno
24	672	49.9	270	20	AAU74136	Human prostate tum
25	600	44.5	204	22	ABG13125	Novel human diagno
26	566.5	42.0	475	22	ABG28310	Novel human diagno
27	566.5	42.0	485	22	ABG13126	Novel human diagno
28	126	9.3	593	24	AAG79911	MECT1. Homo sapie
29	125.5	9.3	650	22	AAU28167	Novel human secret
30	125.5	9.3	657	22	AAU28355	Novel human secret
31	122	9.1	522	21	AAG13630	Arabidopsis thalia
32	122	9.1	539	21	AAG13629	Arabidopsis thalia
33	122	9.1	540	21	AAG13628	Arabidopsis thalia
34	118	8.8	593	22	AAU95646	Human protein sequ
35	118	8.8	804	22	ABG20282	Novel human diagno
36	116	8.6	418	23	ABG66738	Human novel polype
37	115	8.5	471	24	AAO19620	Human nucleic acid
38	114	8.5	1518	24	ABU18375	Breast specific re
39	113.5	8.4	260	21	AAG42023	Arabidopsis thalia
40	113.5	8.4	395	21	AAG42022	Arabidopsis thalia
41	112.5	8.3	532	22	AAU33219	Novel human secret
42	111.5	8.3	1185	20	AAU33497	Human atrophin I p
43	111	8.2	1401	22	ABG24037	Novel human diagno
44	110.5	8.2	778	17	AAU94900	Human tastin. Hom
45	110.5	8.2	778	21	AAU23359	Human tastin prote

ALIGNMENTS

RESULT 1

AAB18449
ID AAB18449 standard; Protein; 252 AA.

XX AAB18449;

AC AAB18449;

XX 15-JAN-2001 (first entry)

DT A human TANGO 261 polypeptide.

DE TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
cellular proliferation; cellular differentiation; cellular adhesion;
von Willebrand factor-associated disorder; cell trafficking; cancer;
hematopoietic associated disease; atelectasis; pulmonary congestion;
oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
intestinal disorder; spleen associated disease; renal disorder;
cardiovascular disorder; ischemic heart disease; hydrocephalus;
brain herniation; iatrogenic disease; inflammation; meningitis;
Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.


OS Homo sapiens.

XX Key Location/Qualifiers
Peptide 1..28
Protein /note= "signal peptide"
29..253
/note= "mature protein"

XX WO200052022-A1.


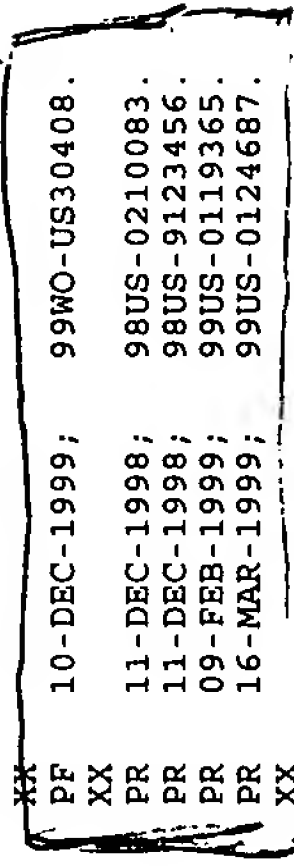
XX 08-SEP-2000.

PD

XX PF 01-MAR-2000; 2000WO-US05226.
XX PR 01-MAR-1999; 99US-0122458. 
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX PT WPI; 2000-579269/54.
XX DR N-PSDB; AAA75151.
XX PT Novel human and murine secreted proteins designated TANGO 216, 261, 262, 266 and 267 useful as modulating agents of cellular processes, e.g. for treating cancer -
XX PS Claim 2; Fig 5; 175pp; English.
XX CC The present sequence represents a human TANGO 261 polypeptide. The specification also describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO polypeptides can be used to modulate cellular proliferation, modulate cellular differentiation and/or modulate cellular adhesion. The proteins can be used to treat any von Willebrand factor-associated disorder, regulate extracellular matrix structuring, cellular adhesion, and cell trafficking and/or migration, modulate cellular interactions, modulate cell adhesion in proliferative disorders, such as cancer, modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, and leukocytes, treat bone marrow, blood and hematopoietic associated diseases and disorders, atelectasis, pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and bronchiectasis, intestinal disorders, spleen associated diseases, modulate renal disorders, treat cardiovascular disorders such as ischemic heart disease, modulate the proliferation, differentiation, and/or function of bone and cartilage cells and to treat bone and/or cartilage associated diseases or ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic disease, inflammations, bacterial and viral meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers, hydrocephalus and encephalitis, and treat hepatic disorders.
XX SQ Sequence 252 AA;
Query Match 100.0%; Score 1348; DB 21; Length 252;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db 1 MAELFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRFHFRFQPTPYLQHEIDLPTISLSDG 120
Db 61 ESTVSGNGIPEQVYAPPRPTDRLAVPPFAQRFHFRFQPTPYLQHEIDLPTISLSDG 120
Qy 121 EEPPTYQGCTQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 180
Db 121 EEPPTYQGCTQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 180
Qy 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 2
AAB01388
ID AAB01388 standard; Protein; 252 AA.
XX

AC AAB01388;
XX 20-OCT-2000 (first entry)
DT Neuron-associated protein.
XX
DE
XX
KW Neuron associated protein; NEUAP; neurological disorder; epilepsy; ischemic cerebrovascular disease; stroke; cerebral neoplasm; Alzheimer's disease; Pick's disease; Huntington's disease; dementia; Parkinson's disease; demyelinating disease; meningitis; prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy; muscular dystrophy; central nervous system; CNS; peripheral nervous system; PNS; myopathy; schizophrenia; actinic keratosis; arteriosclerosis; atherosclerosis; bursitis; cirrhosis; hepatitis; mixed connective tissue disease; MCTD; myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer; autoimmune disease; inflammation; acquired immunodeficiency syndrome; AIDS; Addison's disease; adult respiratory distress syndrome; allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma; Werner syndrome, trauma; human.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..33 /label= Signal peptide
FT Region 9..27 /label= Transmembrane region
FT Modified-site 33 /note= "Potential phosphorylation site"
FT Modified-site 65..68 /note= "Glycosaminoglycan attachment site"
FT Modified-site 81 /note= "Potential phosphorylation site"
FT Modified-site 116 /note= "Potential phosphorylation site"
FT Modified-site 118 /note= "Potential phosphorylation site"
FT Modified-site 147 /note= "Potential phosphorylation site"
FT Modified-site 153 /note= "Potential phosphorylation site"
FT Modified-site 155 /note= "Potential glycosylation site"
FT Modified-site 164 /note= "Potential phosphorylation site"
FT Modified-site 196 /note= "Potential phosphorylation site"
FT Modified-site 219 /note= "Potential phosphorylation site"
FT
XX WO200034477-A2.
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US30408.
XX
PR 11-DEC-1998; 98US-0210083.
PR 11-DEC-1998; 98US-9123456.
PR 09-FEB-1999; 99US-0119365.
PR 16-MAR-1999; 99US-0124687.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-young J, Yang J;
PI Lu DAM, Azimzai Y;
XX
DR WPI; 2000-423423/36.
DR N-PSDB; AAA47429.
XX
PT New human neuron-associated proteins and polynucleotides encoding them, useful for diagnosis, treatment and prevention of cell proliferative disorders including cancer, neuronal and neurological disorders

Claim 1; Page 113-114; 145pp; English.

Human neuron-associated proteins (NEUAP) can be used for for treating or preventing a disorder associated with decreased expression or activity of NEUAP. Antagonists of NEUAP are useful for treating or preventing disorder associated with increased expression or activity of NEUAP. NEUAP or their fragments or derivatives are useful for treating neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia and Parkinson's disease. NEUAPs are also useful for treating other demyelinating diseases, bacterial and viral meningitis, prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and metabolic diseases of the nervous system, neurofibromatosis, other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, inherited, metabolic, endocrine, and toxic myopathies, mental disorders including mood, anxiety and schizophrenic disorders, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal parasitic, protozoal, and helminthic infections, and trauma. This protein was given the Incyte ID no. 1871288CD1.

OS Mammalia.
XX WO200230268-A2.
PN
XX
XX 18-APR-2002.
XX
XX 12-OCT-2001; 2001WO-US32045.
XX
XX 13-OCT-2000; 2000US-0687576.
PR 08-DEC-2000; 2000US-0733288.
PR 08-DEC-2000; 2000US-0733742.
PR 24-JAN-2001; 2001US-263957P.
PR 16-MAR-2001; 2001US-276791P.
PR 16-MAR-2001; 2001US-276888P.
PR 06-APR-2001; 2001US-281922P.
PR 24-APR-2001; 2001US-286214P.
PR 30-APR-2001; 2001US-0847046.
PR 04-MAY-2001; 2001US-288589P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
XX WPI; 2002-471335/50.
DR N-PSDB; ABK92120.
XX
XX Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue -
XX
PS Claim 27; Page 305; 436pp; English.

RESULT 3
ABG61805
ID ABG61805 standard; Protein; 252 AA.
XX
XX ABG61805;
AC
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #6.
XX
XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
KW
XX

Db		241 SKEKDQKGHPL 252	
RESULT 4			
ABP75660			
ID	ABP75660	standard; Protein; 285 AA.	
XX			
AC	ABP75660;		
XX			
DT	10-FEB-2003	(first entry)	
XX			
DE	Human secretory polypeptide SPTM SEQ ID NO 844.		
XX			
KW	Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;		
KW	asthma; Crohn's disease; neurological disorder; epilepsy; cancer;		
KW	Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;		
KW	multiple sclerosis; Parkinson's disease; cell proliferative disorder;		
KW	anti-inflammatory; immunosuppressive; neuroprotective; nootropic;		
KW	neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;		
KW	antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;		
KW	secretory polynucleotide; secretory protein.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200283876-A2.		
XX			
PD	24-OCT-2002.		
XX			
PF	27-MAR-2002; 2002WO-US09921.		
XX			
PR	29-MAR-2001; 2001US-280067P.		
PR	29-MAR-2001; 2001US-280068P.		
PR	16-MAY-2001; 2001US-291280P.		
PR	17-MAY-2001; 2001US-291829P.		
PR	17-MAY-2001; 2001US-291849P.		
PR	19-JUN-2001; 2001US-299428P.		
PR	20-JUN-2001; 2001US-299776P.		
PR	20-JUN-2001; 2001US-300001P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;		
PI	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;		
PI	Daughter SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;		
PI	Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;		
PI	Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;		
XX			
DR	WPI; 2003-075543/07.		
DR	N-PSDB; ABZ36103.		
XX			
PT	New human secretory proteins and polynucleotides, useful for		
PT	diagnosing, treating or preventing autoimmune/inflammatory disorders		
PT	(e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell		
PT	proliferations or cancers -		
XX			
PS	Claim 27; SEQ ID NO 844; 458pp + Sequence Listing; English.		
XX			
CC	The invention relates to a secretory polynucleotide (designated sptm)		
CC	comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a		
CC	naturally occurring polynucleotide sequence at least 90 % identical to		
CC	the polynucleotide sequence, a polynucleotide complementary to them, or an		
CC	RNA equivalent of them. The polypeptide or polynucleotide are useful for		
CC	treating, preventing or diagnosing a disease or condition associated with		
CC	the expression of functional SPTM. These are particularly useful for		
CC	diagnosing, treating or preventing autoimmune/inflammatory disorders		
CC	(e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's		
CC	disease), neurological disorders (e.g. epilepsy, Huntington's disease,		
CC	dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,		
CC	multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,		
CC	schizophrenia or amnesia), or cell proliferative disorders (e.g.		
CC	psoriasis, polycythemia vera, or cancers including adenocarcinoma,		
CC	leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,		
CC			
CC	breast, cervix or prostate). The present sequence is one of the SPTM		
CC	proteins of the invention (ABP75384-ABP75962).		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO at		
CC	ftp.wipo.int/pub/published_pct_sequences.		
XX			
SQ	Sequence	285 AA;	
	Query Match	100.0%; Score 1348; DB 24; Length 285;	
	Best Local Similarity	100.0%; Pred. No. 1.5e-111;	
	Matches 252; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MAELEFVQIIIIIVVMVMVVVITCLLSHYKLSARSFISRHSQRRREDALSSGCLMPS	60
Db	34	MAELEFVQIIIIIVVMVMVVVITCLLSHYKLSARSFISRHSQRRREDALSSGCLMPS	93
Qy	61	ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERPHRFQPTYPYLQHEIDLPTISLSDG	120
Db	94	ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERPHRFQPTYPYLQHEIDLPTISLSDG	153
Qy	121	EEPPPYQGPCITQLRDEQOQLNRESVRAPPNRTIFSDLMDSARLGCGCPPSSNSGIS	180
Db	154	EEPPPYQGPCITQLRDEQOQLNRESVRAPPNRTIFSDLMDSARLGCGCPPSSNSGIS	213
Qy	181	ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW	240
Db	214	ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW	273
Qy	241	SKEKDQKGHPL 252	
Db	274	SKEKDQKGHPL 285	
RESULT 5			
AAB18461			
ID	AAB18461	standard; Protein; 252 AA.	
XX			
AC	AAB18461;		
XX			
DT	15-JAN-2001	(first entry)	
XX			
DE	A human TANGO 261 polypeptide clone.		
XX			
KW	TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;		
KW	cellular proliferation; cellular differentiation; cellular adhesion;		
KW	von Willebrand factor-associated disorder; cell trafficking; cancer;		
KW	hematopoietic associated disease; atelectasis; pulmonary congestion;		
KW	oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;		
KW	intestinal disorder; spleen associated disease; renal disorder;		
KW	cardiovascular disorder; ischemic heart disease; hydrocephalus;		
KW	brain herniation; iatrogenic disease; inflammation; meningitis;		
KW	Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;		
KW	multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.		
OS	Homo sapiens.		
XX			
PN	WO200052022-A1.		
XX			
PD	08-SEP-2000.		
XX			
PF	01-MAR-2000; 2000WO-US05226.		
XX			
PR	01-MAR-1999; 99US-0122458.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Barnes TM, Holtzman DA, Sharp JD, Fraser CC;		
XX			
DR	WPI; 2000-579269/54.		
DR	N-PSDB; AAA75163.		
XX			
PT	Novel human and murine secreted proteins designated TANGO 216, 261,		
PT	262, 266 and 267 useful as modulating agents of cellular processes,		

PT e.g. for treating cancer -
PS Disclosure; Page -; 175pp; English.
XX
CC AAB18461-63 represent human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis, bronchial
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
SQ Sequence 252 AA;
Query Match 99.8%; Score 1345; DB 21; Length 252;
Best Local Similarity 99.6%; Pred. No. 2.3e-111;
Matches 251; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db 1 MADLEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Qy 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180
Db 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180
Qy 181 ATCYSGGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252
RESULT 6
AAB18462
ID AAB18462 standard; Protein; 252 AA.
XX
AC AAB18462;
XX
DT 15-JAN-2001 (first entry)
XX
DE A human TANGO 261 polypeptide clone.
XX
KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;

KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.
XX
OS Homo sapiens.
XX WO200052022-A1.
PN
XX
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX WPI; 2000-579269/54.
DR N-PSDB; AAA75164.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -
XX
PS Disclosure; Page -; 175pp; English.
XX
CC AAB18461-63 represent human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis, bronchial
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
SQ Sequence 252 AA;
Query Match 99.8%; Score 1345; DB 21; Length 252;
Best Local Similarity 99.6%; Pred. No. 2.3e-111;
Matches 251; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Qy 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180
Db 121 EEPYPYQGPCTQLQRPDPEQQLNRESVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGIS 180
Qy 181 ATCYSGGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of
CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity
CC with a candidate agent, where the molecule is a nucleic acid molecule
CC comprising MIVR-1, IEX-1, VDP-1, BTG-2 and TIS-11d or its
CC expression product, determining if the anti-apoptotic activity is
CC modulated and thereby identifying a modulator. The cardiac cell anti-
CC apoptotic molecules and nucleic acids of the invention are useful for
CC treating, diagnosing and monitoring progression of such diseases and
CC disorders as characterised by increased apoptotic cell-death of vascular
CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,
CC stroke, arteriosclerosis and heart failure. The present sequence
CC represents human MIVR-1.

XX
SQ Sequence 287 AA;

Query Match 99.4%; Score 1340; DB 23; Length 287;
Best Local Similarity 99.2%; Pred. No. 7.6e-111;
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSQRRREDALSSEGCLWPS 60
Db : |||||
36 ITELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSQRRREDALSSEGCLWPS 95

Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTTISLDG 120
Db |||||
96 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTTISLDG 155

Qy 121 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGIS 180
Db |||||
156 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGIS 215

Qy 181 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db |||||
216 ATCYSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 275

Qy 241 SKEKDQKQGHPL 252
Db |||||
276 SKEKDQKQGHPL 287

RESULT 9
AAM38712
ID AAM38712 standard; Protein; 237 AA.
XX
AC AAM38712;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 1857.
XX
KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI57868.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PS Example 3; SEQ ID NO 1857; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX SQ Sequence 237 AA;

Query Match 95.1%; Score 1282; DB 22; Length 237;
Best Local Similarity 100.0%; Pred. No. 8.5e-106;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 MMVMVVITCLLSHYKLSARSFISRHSQRRREDALSSEGCLWPSSTVSGNGIPEPQVY 75
Db |||||
1 MMVMVVITCLLSHYKLSARSFISRHSQRRREDALSSEGCLWPSSTVSGNGIPEPQVY 60

Qy 76 APPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTTISLDGEEPPPYQGPTQLQLR 135
Db |||||
61 APPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTTISLDGEEPPPYQGPTQLQLR 120

Qy 136 DPEQQLNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYSGGRMEGPPP 195
Db |||||
121 DPEQQLNRESVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYSGGRMEGPPP 180

Qy 196 TYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKQGHPL 252
Db |||||
181 TYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKQGHPL 237

RESULT 10
AAU78236
ID AAU78236 standard; Protein; 274 AA.
XX
AC AAU78236;
XX
DT 05-JUN-2002 (first entry)
XX
DE Mouse Mechanically Induced Vascular Receptor 1, MIVR-1.
XX
KW Mouse; MIVR-1; Mechanically Induced Vascular Receptor 1;
KW cytostatic; cardiant; cerebroprotective; antiarteriosclerotic;
KW cardiac cell; anti-apoptotic; vascular endothelial cell;
KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;
XX heart failure.
OS Mus musculus.
XX

Db 121 RLGGPCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTR 180
QY 226 LHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 181 LHHTHIAPLESAAIWSKEKDQKGHPL 207

RESULT 12
AAB18450
ID AAB18450 standard; Protein; 217 AA.

XX AAB18450;

DT 15-JAN-2001 (first entry)

DE A murine TANGO 261 polypeptide.

XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

XX Mus sp.

XX WO200052022-A1.

PN 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US05226.

XX 01-MAR-1999; 99US-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX WPI; 2000-579269/54.

DR N-PSDB; AAA75152.

XX Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -

XX Claim 8; Fig 6A-B; 175pp; English.

PS The present sequence represents a murine TANGO 261 polypeptide. The
XX specification also describes TANGO 266, TANGO 216, TANGO 262, and
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular
CC proliferation, modulate cellular differentiation and/or modulate
CC cellular adhesion. The proteins can be used to treat any von Willebrand
CC factor-associated disorder, regulate extracellular matrix structuring,
CC cellular adhesion, and cell trafficking and/or migration, modulate
CC cellular interactions, modulate cell adhesion in proliferative
CC disorders, such as cancer, modulate the proliferation, differentiation,
CC and/or function of cells that appear in the bone marrow, and leukocytes,
CC treat bone marrow, blood and hematopoietic associated diseases and
CC disorders, atelectasis, pulmonary congestion or oedema, emphysema,
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal
CC disorders, spleen associated diseases, modulate renal disorders, treat
CC cardiovascular disorders such as ischemic heart disease, modulate the
CC proliferation, differentiation, and/or function of bone and cartilage
CC cells and to treat bone and/or cartilage associated diseases or
CC disorder. They may also be used to treat disorders associated with the
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic
CC disease, inflammations, bacterial and viral meningitis, Alzheimer's
CC Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,

CC brain cancers, hydrocephalus and encephalitis, and treat hepatic
CC disorders.
XX Sequence 217 AA;

QY Query Match 79.1%; Score 1066.5; DB 21; Length 217;
Best Local Similarity 88.9%; Pred. No. 1.le-86;
Matches 201; Conservative 7; Mismatches 9; Indels 9; Gaps 3;

QY 27 LSHYKLSARSFISRHSGRRREDALSSGCLWPSESTVSGNGIPEQVYAPRPTDRLAV 86
Db 1 LSHYKLSARSFISRHSGARRDDGLSSEGCLWPSESTVSG-GMPEQVYAPRPTDRLAV 59

QY 87 PPFQQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPYQGPCTQLQRDPEQQLELNRE 146
Db 60 PPFIQRS---RFQPTYPYLQHEIALPTTISLSDGEEPPYQGPCTQLQRDPEQQLELNRE 116

QY 147 SVRAPPNRTIFDSDLMDSARLGGPCPPSSNSGISATCYSGGRMEGPPPTYSEVIGHYPG 206
Db 117 SVRAPPNRTIFDSDLIDSTMLGGPCPPSSNSGISATCYSGGRMEGPPPTYSEVIGHYPG 176

QY 207 SSFQHQSSGPPSSLLEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 177 SSFQHQSSNGPSSLLEGTRLHSHIAPLE-----NKEKEKQKGHPL 217

RESULT 13
ABG77052

ID ABG77052 standard; Protein; 241 AA.

XX ABG77052;

XX 05-NOV-2002 (first entry)

DT Prostate specific protein #22.

XX Human; prostate specific nucleic acid; PSNA; prostate cancer;
KW cytostatic; non-cancerous prostate disease; PSP; vaccine;
KW prostate specific protein; metastasis.

XX Homo sapiens.

XX WO200242776-A2.

XX 30-MAY-2002.

XX 01-NOV-2001; 2001WO-US45654.

XX 01-NOV-2000; 2000US-244782P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Recipon H, Chen S, Liu C;

XX WPI; 2002-490217/52.

XX New polypeptide useful for diagnosing and monitoring the presence and
PT metastases of prostate cancer in a patient and as a component in
PT databases for search analysis as well as in sequence analysis
PT algorithms -

XX Claim1; Page 213-214; 242pp; English.

CC The invention relates to an isolated polypeptide comprising a sequence
CC with 60 % identity to one of prostate specific protein (PSP) sequences,
CC or comprising an amino acid sequence encoded by one of 136 nucleotide
CC prostate specific nucleic acids sequences, PSNA, (or a sequence that
CC hybridises to it or is 60% identical to it), given in the specification.
CC Also included are a vector comprising the polynucleotide, a host cell
CC comprising the vector, an antibody specific for the PSP proteins and a
CC vaccine comprising the protein or polynucleotide. The PSP and PSNA are
CC useful for diagnosing and monitoring the presence and metastases of
CC prostate cancer in a patient. The PSNA is useful for determining the

CC the level PSNA in a sample. An antibody to the PSP is useful for
CC determining the presence of prostate specific protein in a sample, and
CC for treating a patient with prostate cancer, which induces an immune
CC response against the prostate cancer cell expressing the nucleic acid or
CC polypeptide and a kit is useful for detecting a risk of cancer or
CC presence of cancer in a patient. PSNA is useful as hybridisation probes
CC to detect, characterise and quantify hybridising nucleic acids from both
CC genomic and transcript-derived nucleic acid samples and also in
CC microarrays. Sequences of PSP and PSNA are useful as components in
CC databases for search analysis as well as in sequence analysis algorithms.
CC PSNA is useful to drive in vivo expression of PSP. The present
CC sequence is a PSP of the invention.

XX Sequence 241 AA;

Query Match 79.0%; Score 1065.5; DB 23; Length 241;
Best Local Similarity 82.9%; Pred. No. 1.5e-86;
Matches 199; Conservative 0; Mismatches 0; Indels 41; Gaps 1;

Qy 54 EGCLWPSESTVSGNGIPE-----P 72
Db 1 EGCLWPSESTVSGNGIPECCPCWDPPCRRSSAPCPAGSSPALCSLHTGARTLPLFGGGRP 60
Qy 73 QVYAPRPTDRLAVPPFAQRERFHFQPTYPYLQHEIDLPTISLSDGEEPPPYQGCTL 132
Db 61 QVYAPRPTDRLAVPPFAQRERFHFQPTYPYLQHEIDLPTISLSDGEEPPPYQGCTL 120
Qy 133 QLRDPEQQLNRESVRAPPNRTIFDSDLMDARSALGGPCPPSSNSGISATCYGSGGRMEG 192
Db 121 QLRDPEQQLNRESVRAPPNRTIFDSDLMDARSALGGPCPPSSNSGISATCYGSGGRMEG 180
Qy 193 PPPTYSEVIGHYGGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 181 PPPTYSEVIGHYGGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 240

RESULT 14
AAB18464
ID AAB18464 standard; Protein; 217 AA.

XX AC AAB18464;

DT 15-JAN-2001 (first entry)

DE A murine TANGO 261 polypeptide clone.

XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

XX OS Mus sp.

XX PN WO200052022-A1.

XX PD 08-SEP-2000.

XX PF 01-MAR-2000; 2000WO-US05226.

XX PR 01-MAR-1999; 99US-0122458.

XX PA (MILL-) MILLENNIUM PHARM INC.

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX WPI; 2000-579269/54.

DR N-PSDB; AAB75166.

XX

PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -

PS Disclosure; Page -; 175pp; English.

XX AAB18464-66 represent murine TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis, bronchial
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.

XX Sequence 217 AA;

Query Match 78.9%; Score 1063.5; DB 21; Length 217;
Best Local Similarity 88.5%; Pred. No. 1.9e-86;
Matches 200; Conservative 8; Mismatches 9; Indels 9; Gaps 3;

Qy 27 LSHYKLSARSFISRHSGRRREDALSSGCLWPSESTVSGNGIPEQVYAPRPTDRLAV 86
Db 1 LSHYKLSARSFISRHSGRRRDDGLSSDGLWPSESTVSG-GMPEQVYAPRPTDRLAV 59

Qy 87 PPFAQRERFHFQPTYPYLQHEIDLPTISLSDGEEPPPYQGCTLQLRDPEQQLNRE 146
Db 60 PPFIQRS---RFQPTYPYLQHEIALPTISLSDGEEPPPYQGCTLQLRDPEQQLNRE 116

Qy 147 SVRAPPNRTIFDSDLMDARSALGGPCPPSSNSGISATCYGSGGRMEGPPPTYSEVIGHYPG 206
Db 117 SVRAPPNRTIFDSDLMDSTMLGGPCPPSSNSGISATCYGSGGRMEGPPPTYSEVIGHYPG 176

Qy 207 SSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 177 SSFQHQSSGPPSLLEGTRLHHTHIAPLE-----NKEKEKQKGHPL 217

RESULT 15

AAB18465

ID AAB18465 standard; Protein; 217 AA.

XX AC AAB18465;

XX DT 15-JAN-2001 (first entry)

XX DE A murine TANGO 261 polypeptide clone.

XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;

KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
XX multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder.

OS Mus sp.

XX WO200052022-A1.

PN XX

XX PD

XX 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US05226.

XX 01-MAR-1999; 99US-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX WPI; 2000-579269/54.

XX N-PSDB; AAA75167.

XX Novel human and murine secreted proteins designated TANGO 216, 261,

XX 262, 266 and 267 useful as modulating agents of cellular processes,

XX e.g. for treating cancer -

XX Disclosure; Page -; 175pp; English.

CC AAB18464-66 represent murine TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.

XX Sequence 217 AA;

Query Match 78.9%; Score 1063.5; DB 21; Length 217;
Best Local Similarity 88.5%; Pred. No. 1.9e-86;
Matches 200; Conservative 8; Mismatches 9; Indels 9; Gaps 3;
Qy 27 LSHYKLSARSFISRHSQRRRDGLSSEGCLWPSSTVSGNGIPEPQVYAPRPTDRLAV 86
Db 1 LSHYKLSARSFISRHSQRRRDGLSSEGCLWPSSTVSG-GMPEPQVYAPRPTDRLAV 59
Qy 87 PPFQQRERFRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGPCTQLQRPDPEQQLELNRE 146
Db 60 PPFQQR--RFQPTYPYLQHEIALPPTISLSDGEEPPPYQGPCTQLQRPDPEQQLELNRE 116
Qy 147 SVRAPPNRTIFDSDLMDSARLGPCPPSSNSGISATCYSGGRMEGPPPTTYSEVIGHYPG 206
Db 117 SVRAPPNRTIFDSDLIDSTMLGGPCPPSSNSGISATCYSSGRMEGPPPTTYSEVIGHYPG 176
Qy 207 SSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKGHPL 252
Db 177 SSFQHQSSNGPSSLLEGTRLHHTHIAPLE-----NKEKEKQKGHPL 217

THIS PAGE BLANK (U8PT0)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 8, 2003, 03:11:01 ; Search time 4129 Seconds
(without alignments)
10512.258 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tcctccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_nam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	953	89.8	1061	9	BC015918	BC015918 Homo sapi
2	926	87.3	1141	9	AF224278	AF224278 Homo sapi
3	903	85.1	1818	9	AY128643	AY128643 Homo sapi
4	902	85.0	4839	9	AF305616	AF305616 Homo sapi
5	749	70.6	1321	6	AX392417	AX392417 Sequence
6	701	66.1	861	6	AX392419	AX392419 Sequence
7	695	65.5	61505	9	AF305426	AF305426 Homo sapi
8	695	65.5	130435	9	HS718J7	AL035541 Human DNA
9	509	48.0	1583	6	AX593655	AX593655 Sequence
10	350	33.0	408	6	AX071267	AX071267 Sequence
11	322	30.3	812	6	AX011709	AX011709 Sequence
12	313	29.5	693	6	AX392430	AX392430 Sequence
13	157	14.8	150224	9	HSJ1059L7	AL121913 Human DNA
14	79	7.4	2570	9	AK056098	AK056098 Homo sapi
15	51	4.8	51	6	AX199565	AX199565 Sequence
16	44	4.1	651	10	AF220208	AF220208 Mus muscu
17	44	4.1	878	6	AX392428	AX392428 Sequence
18	44	4.1	1379	10	BC036995	BC036995 Mus muscu
19	44	4.1	156698	10	AL837509	AL837509 Mouse DNA
20	44	4.1	175754	2	AC110189	AC110189 Mus muscu
21	44	4.1	176458	2	AC134911	AC134911 Mus muscu
22	44	4.1	176821	2	AL837520	AL837520 Mus muscu
23	42	4.0	673	6	AX525744	AX525744 Sequence
24	42	4.0	249554	2	AC139417	AC139417 Rattus no
25	42	4.0	258632	2	AC111878	AC111878 Rattus no
26	32	3.0	756	10	AB045588S2	AB045589 Rattus no
27	32	3.0	56475	9	AL353658	AL353658 Human DNA
28	32	3.0	110000	2	AC109084_1	Continuation (2 of
29	32	3.0	161462	2	AC136669	AC136669 Rattus no
30	32	3.0	172525	9	AC008514	AC008514 Homo sapi
31	32	3.0	179372	2	AC040922	AC040922 Homo sapi
32	32	3.0	190025	2	AC140145	AC140145 Homo sapi
33	32	3.0	190149	2	AC122097	AC122097 Rattus no
34	32	3.0	192183	2	AC134492	AC134492 Rattus no
35	32	3.0	198582	9	AC005291	AC005291 Homo sapi
36	32	3.0	220084	2	AC123651	AC123651 Mus muscu
37	32	3.0	221080	2	AC106160	AC106160 Rattus no
38	32	3.0	226791	2	AC094199	AC094199 Rattus no
39	32	3.0	239018	2	AC106202	AC106202 Rattus no
40	32	3.0	242543	2	AC096128	AC096128 Rattus no
41	32	3.0	249703	2	AC126988	AC126988 Rattus no
42	32	3.0	258545	2	AC135666	AC135666 Rattus no
43	32	3.0	259335	2	AC126194	AC126194 Rattus no
44	32	3.0	279534	2	AC114035	AC114035 Rattus no
45	32	3.0	281747	2	AC094865	AC094865 Rattus no

ALIGNMENTS

RESULT 1
BC015918
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BC015918 1061 bp mRNA linear PRI 17-OCT-2001
Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.

BC015918
MGC:
BC015918.1 GI:16198474

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1061)

Strausberg,R.
Direct Submission
Submitted (15-OCT-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ruben Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: e Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

source

1. .1061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20374 IMAGE:4559576"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC 14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
131. .844
/codon_start=1
/product="Unknown (protein for MGC:20374)"
/protein_id="AAH15918.1"
/db_xref="GI:16198475"
/translation="MMVMVVVITCLLSHYKLSRFSFISRHSGRRRREDALSSGCLWP
SESTVSGNGIPEPQVYAPRPTDR LAVPPFAQREFRFPQTPYLYQHEIDLPTISL
SDGEPPPYQGPTQLRDPQQLNRESVRAPNRTIFDSDLMD SARLGGPCPPSS
NSGISATCYGSGRMGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAP
LESAAIWSKEKDQKHPL"

BASE COUNT 239 a 350 c 311 g 161 t

ORIGIN

Query Match 89.8%; Score 953; DB 9; Length 1061;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	105	CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGGTGGTGG 164
Db	90	CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGGTGGTGG 149
Qy	165	TGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA 224
Db	150	TGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA 209
Qy	225	GCCAGGGCGGAGGAGAGATGCCCTGTCTCCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 284
Db	210	GCCAGGGCGGAGGAGAGATGCCCTGTCTCCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 269
Qy	285	GCACAGTGTCAAGCAACGGAATCCAGAGCCGCAAGTCTACGCCCGCCCTCGGCCACCG 344
Db	270	GCACAGTGTCAAGCAACGGAATCCAGAGCCGCAAGTCTACGCCCGCCCTCGGCCACCG 329
Qy	345	ACCGCCTGGCGTGC CGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACT 404
Db	330	ACCGCCTGGCGTGC CGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACT 389

Qy	405	ATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCACCATCTCGTGTTCAGACGGGAGG 464
Db	390	ATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCACCATCTCGTGTTCAGACGGGAGG 449
Qy	465	AGCCCCACCTACAGGGCCCCCTGCACCTCCAGTTCGGGACCCCGAGCAGCAGCTGG 524
Db	450	AGCCCCACCTACAGGGCCCCCTGCACCTCCAGTTCGGGACCCCGAGCAGCAGCTGG 509
Qy	525	AACTGAACCGGGAGTCCGTGCGCGACCCCTGCCGCCACCACCAACAGAACCATCTTCGACAGTACCTGA 584
Db	510	AACTGAACCGGGAGTCCGTGCGCGACCCCTGCCGCCACCACCAACAGAACCATCTTCGACAGTACCTGA 569
Qy	585	TGGATAGTCCAGGCTGGCGGGCCCTGCCGCCACCACCAACAGAACCATCTTCGAGCATCAGCGCCA 644
Db	570	TGGATAGTCCAGGCTGGCGGGCCCTGCCGCCACCACCAACAGAACCATCTTCGAGCATCAGCGCCA 629
Qy	645	CGTGCTACGGCAGCGGGCGGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTCTATCG 704
Db	630	CGTGCTACGGCAGCGGGCGGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTCTATCG 689
Qy	705	GCCACTACCCGGGTCTCTCTCCAGCACGAGCAGTGGGCGCCCTCCTTGTCTGG 764
Db	690	GCCACTACCCGGGTCTCTCTCCAGCACGAGCAGTGGGCGCCCTCCTTGTCTGG 749
Qy	765	AGGGGACCCGGTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 824
Db	750	AGGGGACCCGGTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 809
Qy	825	AAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCGGGGGCGGGCTGGGG 884
Db	810	AAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCGGGGGCGGGCTGGGG 869
Qy	885	CTGCGTAGGTGAAAAGGCAGAACACTCCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCG 944
Db	870	CTGCGTAGGTGAAAAGGCAGAACACTCCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCG 929
Qy	945	GGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA 1004
Db	930	GGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA 989
Qy	1005	CATGTGATGTCTGTCTGAATGCACAAGTAAAGAGAGCTTGCAAAAAA 1057
Db	990	CATGTGATGTCTGTCTGAATGCACAAGTAAAGAGAGCTTGCAAAAAA 1042

RESULT 2

AF224278

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Xu, L.L., Shanmugam, N., Segawa, T., Sesterhenn, I.A., McLeod, D.G.,

Moul, J.W. and Srivastava, S.

A novel androgen-regulated gene, PMEPAL, located on chromosome

20q13 exhibits high level expression in prostate

Genomics 66 (3), 257-263 (2000)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Xu, L.L., Shanmugam, N., Segawa, T., Sesterhenn, I.A., McLeod, D.G.,

Moul, J.W. and Srivastava, S.

Direct Submission

TITLE

JOURNAL

Submitted (12-JAN-2000) CPDR, USUHS, 1530 East Jefferson Street,

Rockville, MD 20852, USA

FEATURES

Location/Qualifiers

source

1. .1141

		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/chromosome="20"	
		/map="20q13.31-13.33"	
		/cell_line="LNCap"	
gene		1..1141	
		/gene="PMEPA1"	
CDS		96..854	
		/gene="PMEPA1"	
		/note="type 1b transmembrane protein; expression is induced in response to the synthetic androgen, R1881; expression is abundant in, and restricted to, prostate glandular epithelial cells; similar to the predicted protein encoded by sequence deposited at GenBank Accession Number NP_004329"	
		/codon_start=1	
		/product="PMEPA1 protein"	
		/protein_id="AAF86322.1"	
		/db_xref="GI:9255809"	
		/translation="MAELEFVQIIIVVMVMVVITCLLSHYKLSARSFIRHSQRRREDALSSGCLWPESESTVSGNGIPEPQVYAPRPTRDLAVPPFAQRERFHRFQPTYPLQHEIDLPTISLSDGEEPPPYQGPTLQLRDEPQQLNRESVRAPNRTIFDSDLMSARLGGPCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDQKGHPL"	
BASE COUNT	271 a	350 c	336 g 184 t
ORIGIN			
Query Match	87.3%; Score 926; DB 9; Length 1141;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches	976; Conservative	0; Mismatches	1; Indels 0; Gaps 0;
QY	81	GTCTCCTCGAAACCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGG	140
Db	76	GTCTCCTCGAAACCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGG	135
QY	141	TGGTGATGATGGTGATGGTGATGATCAGTGCCCTGCTGAGCCACTACAAGCTGTCTG	200
Db	136	TGGTGATGATGGTGATGGTGATGATCAGTGCCCTGCTGAGCCACTACAAGCTGTCTG	195
QY	201	CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAG	260
Db	196	CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAG	255
QY	261	AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTGAGGCAACGGAATCCAGAGCCGAGG	320
Db	256	AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTGAGGCAACGGAATCCAGAGCCGAGG	315
QY	321	TCTACGCCCGCCTCGGCCACCGACCGCCTGGCCGTGCCCGCTTCGCCAGCGGGAGC	380
Db	316	TCTACGCCCGCCTCGGCCACCGACCGCCTGGCCGTGCCCGCTTCGCCAGCGGGAGC	375
QY	381	GCTTCACCGCTTCAGCCGCCACCTATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCA	440
Db	376	GCTTCACCGCTTCAGCCGCCACCTATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCA	435
QY	441	CCATCTCGTGTGACACGGGGAGGAGCCCCACCCTACCGGGCCCTGCACCTCCAGC	500
Db	436	CCATCTCGTGTGACACGGGGAGGAGCCCCACCCTACCGGGCCCTGCACCTCCAGC	495
QY	501	TTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTCGGTGGCGGCACCCCAACA	560
Db	496	TTCGGGACCCCGAGCAGCTGGAACTGAACCGGGAGTCGGTGGCGGCACCCCAACA	555
QY	561	GAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGGGCCCTGCCCCCCCA	620
Db	556	GAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGGGCCCTGCCCCCCCA	615
QY	621	GCAGTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGGGCGGCATGGAGGGGCCGC	680
Db	616	GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGCCGC	675
QY	681	CGCCACCTACAGCGAGGTATCGGCCACTACCGGGGGTCTCTCTCCAGCACAGCAGAGA	740

Db	676	CGCCACCTACAGGAGGTATCGCCACTACCCGGGTCTCTTCCAGCACCGACAGA	735
Qy	741	GCAGTGGGCGCCCTCTTGCTGGAGGGGACCCGGTCCACACACATCGGCGCCC	800
Db	736	GCAGTGGGCGCCCTCTTGCTGGAGGGGACCCGGTCCACACACATCGGCGCCC	795
Qy	801	TAGAGAGCGCAGCCATCTGGAGCAAAGAGAGAGGATAAAACAGAAAGGACACCCCTCTTAGG	860
Db	796	TAGAGAGCGCAGCCATCTGGAGCAAAGAGAGAGGATAAAACAGAAAGGACACCCCTCTTAGG	855
Qy	861	GTCCCCAGGGGGCGGCTGGGCTGCGTAGGTGAAAGGCAGAACACTCCGCGCTTCT	920
Db	856	GTCCCCAGGGGGCGGCTGGGCTGCGTAGGTGAAAGGCAGAACACTCCGCGCTTCT	915
Qy	921	TAGAAGAGGAGTGAGAGGAAGCGGGGGCGGCGCAGCAACGCATCGTGTGGCCCTCCCCTCC	980
Db	916	TAGAAGAGGAGTGAGAGGAAGCGGGGGCGGCGCAGCAACGCATCGTGTGGCCCTCCCCTCC	975
Qy	981	CACCTCCCTGTGTATAATATTACATGTGATGTCTGCTGAATGCACAAGCTAAGAGA	1040
Db	976	CACCTCCCTGTGTATAATATTACATGTGATGTCTGCTGAATGCACAAGCTAAGAGA	1035
Qy	1041	GCTTGCAAAAAA	1057
Db	1036	GCTTGCAAAAAA	1052
RESULT 3			
AY128643			
LOCUS			
DEFINITION Homo sapiens PMEPA1 variant A protein mRNA, complete cds.			
ACCESSION AY128643			
VERSION AY128643.1 GI:22121998			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE 1 (bases 1 to 1818)			
AUTHORS Brunschwig,E.B., Wilson,K., Mack,D., Dawson,D., Lawrence,E., Willson,J.K., Lu,S., Nosrati,A., Rerko,R.M., Swinler,S., Beard,L., Lutterbaugh,J.D., Willis,J., Platzer,P. and Markowitz,S.			
TITLE PMEPA1, a transforming growth factor-beta-induced marker of terminal colonocyte differentiation whose expression is maintained in primary and metastatic colon cancer			
JOURNAL Cancer Res. 63 (7), 1568-1575 (2003)			
MEDLINE 22557253			
PUBMED 12670906			
REFERENCE 2 (bases 1 to 1818)			
AUTHORS Brunschwig,E.B., Wilson,K., Mack,D., Dawson,D., Lawrence,E., Willson,J.K.V., Lu,S., Nosrati,A., Swinler,S., Beard,L., Lutterbaugh,J.D., Willis,J., Platzer,P. and Markowitz,S.			
TITLE Direct Submission			
JOURNAL Submitted (03-JUL-2002) Department of Medicine, Case Western Reserve University/Howard Hughes Medical Institute, 11001 Cedar Ave., Cleveland, OH 44106, USA			
FEATURES			
source			
1..1818			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="20"			
/map="20q13.31-13.33"			
145..858			
/codon_start=1			
/product="PMEPA1 variant A protein"			
/protein_id="AA089277.1"			
/db_xref="GI:22121998"			
/translation="MMVMVVITCLLSHYKLSARSFIRHSQRRREDALSSGCLWPESESTVSGNGIPEPQVYAPRPTRDLAVPPFAQRERFHRFQPTYSDGEEPPPYQGPTLQLRDEPQQLNRESVRAPNRTIFDSDLMSARLGGPCPPSSNSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAP			
CDS			

BASE COUNT		459 a	467 c	494 g	398 t
ORIGIN					
Query Match 85.1%; Score 903; DB 9; Length 1818;					
Best Local Similarity 99.9%; Pred. No. 0;					
Matches 953; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	104	GC	GAGCTGGAGTTT	GTTCAGATCATCATCGTGGTGGTGATGATGTTGGTG	163
Db	103	GC	GAGCTGGAGTTT	GTTCAGATCATCATCGTGGTGGTGATGATGTTGGTG	162
QY	164	GT	GATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCAC	223	
Db	163	GT	GATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCAC	222	
QY	224	AG	CCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAGATGCTGTGGCCCTCGGAG	283	
Db	223	AG	CCAGGGCGGAGGAGAGATGCCCTGTCTCAGAAGATGCTGTGGCCCTCGGAG	282	
QY	284	AG	CAGTGTGAGGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCCTCGGCCACC	343	
Db	283	AG	CAGTGTGAGGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCCTCGGCCACC	342	
QY	344	GAC	CGCTGGCGCTGCCCTTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACC	403	
Db	343	GAC	CGCTGGCGCTGCCCTTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACC	402	
QY	404	TAT	CCGTACTGCAGCAGCAGATCGACCTGCCGCCACCAATCTCGCTGTCAGACGGGAG	463	
Db	403	TAT	CCGTACTGCAGCAGCAGATCGACCTGCCGCCACCAATCTCGCTGTCAGACGGGAG	462	
QY	464	GAG	CCCCACCTACAGGGGCCCTGGACCTTCCAGCTTCGGGACCCCGAGCAGAGCTG	523	
Db	463	GAG	CCCCACCTACAGGGGCCCTGGACCTTCCAGCTTCGGGACCCCGAGCAGAGCTG	522	
QY	524	GAA	CTGAACGGGAGTCGGTGC	CGCACCCCAACAGAACCATCTTCGACAGTGACCTG	583
Db	523	GAA	CTGAACGGGAGTCGGTGC	CGCACCCCAACAGAACCATCTTCGACAGTGACCTG	582
QY	584	AT	GATAGTCCAGGCTGGCGGCCCTGCCCCCAGCAGTAACCTCGGCGATCAGCGCC	643	
Db	583	AT	GATAGTCCAGGCTGGCGGCCCTGCCCCCAGCAGTAACCTCGGCGATCAGCGCC	642	
QY	644	AC	GTCTACGGCAGCGCGCGCATGGAGGGCGCGCCGCCACCTACAGCAGGTCTATC	703	
Db	643	AC	GTCTACGGCAGCGCGCGCATGGAGGGCGCGCCGCCACCTACAGCAGGTCTATC	702	
QY	704	GG	CACTACCGGGGTCTCTCTCCAGCACAGCAGAGCAGTGGCGCCCTCCTTGCTG	763	
Db	703	GG	CACTACCGGGGTCTCTCTCTCCAGCACAGCAGAGCAGTGGCGCCCTCCTTGCTG	762	
QY	764	GAG	GGACCCGGCTCCACCACACATCGCGCCCTTAGAGCGCAGCCATCTGGAGC	823	
Db	763	GAG	GGACCCGGCTCCACCACACATCGCGCCCTTAGAGCGCAGCCATCTGGAGC	822	
QY	824	AA	GAGAAGGATAAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGGGCTGGG	883	
Db	823	AA	GAGAAGGATAAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGGGCTGGG	882	
QY	884	GCT	CGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGGC	943	
Db	883	GCT	CGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGGC	942	
QY	944	GG	GGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTT	1003	
Db	943	GG	GGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTT	1002	
QY	1004	AC	ATGTATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGC	AAAAA	1057
Db	1003	AC	ATGTATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGC	AAAAA	1056

RESULT 4	AF305616	4839 bp	mRNA	linear	PRI 21-OCT-2001
LOCUS	Homo sapiens STAG1/PMEPA1 mRNA, complete cds.				
DEFINITION	Homo sapiens STAG1/PMEPA1 mRNA, complete cds.				
ACCESSION	AF305616				
VERSION	AF305616.1 GI:16303741				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 4839)				
AUTHORS	Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.				
TITLE	Characterization of a novel gene, STAG1/PMEPA1, upregulated in renal cell carcinoma and other solid tumors				
JOURNAL	Mol. Carcinog. 32 (1), 44-53 (2001)				
MEDLINE	21453682				
PUBMED	11568975				
REFERENCE	2 (bases 1 to 4839)				
AUTHORS	Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-SEP-2000) Centre for Molecular Biotechnology, Queensland University of Technology, 2 George St, Brisbane, QLD 4001, Australia				
FEATURES	Location/Qualifiers				
source	1. .4839				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="20"				
	/map="20q13.2-q13.33"				
	321. .1184				
	/note="unknown function"				
	/codon_start=1				
	/product="STAG1/PMEPA1"				
	/protein_id="AAL16781.1"				
	/db_xref="GI:16303742"				
	/translation="MRLMGVNSTAAAGQPNVSTCNCKRSLFQSMITELEFVQI IIVVMVMVIVITCLLSHYKLSARFISRHSGRRREDALSSGCLWPSESTVSGN GIPEQVYAPRPTDRLAVPPFAQRERFHRFQTPYLYQHEIDLPTISLSDGEEPPP YQGPCTLQLRDPQQLNRESVRAPNRTIFDSDLMDSARLGGPCPPSSNSGISATC YGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHHTIAPLESAAIWS KEKDKQKGHPL"				
	polyA_signal 2158. .2163				
	polyA_signal 2463. .2468				
	polyA_signal 4818. .4823				
BASE COUNT	1103 a	1313 c	1290 g	1133 t	
ORIGIN					
Query Match	85.0%; Score 902; DB 9; Length 4839;				
Best Local Similarity	99.9%; Pred. No. 0;				
Matches 952; Conservative	0;	Mismatches	1;	Indels	0; Gaps 0;
QY	105	CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGTTGGTGG	164		
Db	430	CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGTTGGTGG	489		
QY	165	TGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCACA	224		
Db	490	TGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCACA	549		
QY	225	GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA	284		
Db	550	GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA	609		
QY	285	GCACAGTGTACGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCCACCG	344		
Db	610	GCACAGTGTACGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCCACCG	669		
QY	345	ACCGCCTGGCCGTGCCGCCCTTCGCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT	404		
Db	670	ACCGCCTGGCCGTGCCGCCCTTCGCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT	729		

Qy	405	ATCCGTACCTGCAGCACGAGATCGACCTGCGCGCCACCATCTCGTGT	CAGCGGGAGG	464
Db	730	ATCCGTACCTGCAGCACGAGATCGACCTGCGCACCCACCATCTCGTGT	CAGCGGGAGG	789
Qy	465	AGCCCCACCCCTACCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCGAGCAGCTGG	524	
Db	790	AGCCCCACCCCTACCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCGAGCAGCTGG	849	
Qy	525	AACTGAACGGGAGTCGGTGCAGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGA	584	
Db	850	AACTGAACGGGAGTCGGTGCAGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGA	909	
Qy	585	TGGATAGTGCCAGGCTGGCGGCCCCCTGCCCCCCCCCAGAGTAACTCGGGCATCAGCGCCA	644	
Db	910	TGGATAGTGCCAGGCTGGCGGCCCCCTGCCCCCCCCCAGAGTAACTCGGGCATCAGCGCCA	969	
Qy	645	CGTGACGCGCAGCGCGGCGCATGGAGGGCGCGCCGACCATCTTCGACAGGATCATCG	704	
Db	970	CGTGACGCGCAGCGCGGCGCATGGAGGGCGCGCCGACCATCTTCGACAGGATCATCG	1029	
Qy	705	GCCACTACCCGGGTCTCTCCAGCACCATCGCGCCCTAGAGAGTGCGCGCCCTCTTGCTGG	764	
Db	1030	GCCACTACCCGGGTCTCTCTCCAGCACCATCGCGCCCTAGAGAGTGCGCGCCCTCTTGCTGG	1089	
Qy	765	AGGGACCCCGCTCCACCACACACATCGCGCCCTAGAGAGCGCGCAGCCATCTGGAGCA	824	
Db	1090	AGGGACCCCGCTCCACCACACACATCGCGCCCTAGAGAGCGCGCAGCCATCTGGAGCA	1149	
Qy	825	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCGCGGTGGG	884	
Db	1150	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCGCGGTGGG	1209	
Qy	885	CTGCGTAGGTGAAAAGGCAGAACCTCCGCGCTTCTTAGAGAGGAGTGAGAGGAGCG	944	
Db	1210	CTGCGTAGGTGAAAAGGCAGAACCTCCGCGCTTCTTAGAGAGGAGTGAGAGGAGCG	1269	
Qy	945	GGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA	1004	
Db	1270	GGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATTTA	1329	
Qy	1005	CATGTATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1057	
Db	1330	CATGTATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1382	

RESULT 5
AX392417
LOCUS AX392417 1321 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 1 from Patent WO0216416.
ACCESSION AX392417
VERSION AX392417.1 GI:19700732
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and Turi,T.G.
TITLE Diagnosis and treatment of cardiovascular conditions
JOURNAL Patent: WO 0216416-A 1 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source .
1. .1321
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
413. .1276
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD29005.1"
/db_xref="GI:19700732"
/translation="MHRMLGVNSTAAAGQPNVSVCTCNCKRSLFQSMETELFEVQI

BASE COUNT	223 a	493 c	440 g	165 t
ORIGIN	IIIVVMVMVVITCLLSHYKLSARSFISRHSQRRREDALSSSEGCLWPSESTVSGN GIPEPQVYAPRPTDRLAAPPFAQRERFHRQPTYPYLQHEIDLPTTISLSDGEPPPP YQGPCTLOLRDPEOQLELNRESVRAPNRTTIFDSDLMSARLGGPCPPSSNSGTSATC YSGGRMEGPPPTPYSEVIGHYPGSSFFHQQSSGPPSLLEGTRLHHTHIAPLESAIWS KEKDQKQGHPL"			
Query Match	70.6%; Score 749; DB 6; Length 1321;			
Best Local Similarity	99.9%; Pred. No. 0;			
Matches 799; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
Qy	105	CGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGGTGGTGG	164	
Db	522	CGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGGTGGTGG	581	
Qy	165	TGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA	224	
Db	582	TGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA	641	
Qy	225	GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGATGCCTGTGGCCCTCGGAGA	284	
Db	642	GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGATGCCTGTGGCCCTCGGAGA	701	
Qy	285	GCACAGTGTACGGCAACGGAATCCCAGAGCCGAGGTCTACGCCCCCGCTCGGCCACCG	344	
Db	702	GCACAGTGTACGGCAACGGAATCCCAGAGCCGAGGTCTACGCCCCCGCTCGGCCACCG	761	
Qy	345	ACCGCTGGCCGTGCCCGCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT	404	
Db	762	ACCGCTGGCCGTGCCCGCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCT	821	
Qy	405	ATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGT	CAGACGGGAGG	464
Db	822	ATCCGTACCTGCAGCACGAGATCGACCTGCCACCACCATCTCGTGT	CAGACGGGAGG	881
Qy	465	AGCCCCACCTACCAGGGCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG	524	
Db	882	AGCCCCACCTACCAGGGCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG	941	
Qy	525	AACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCATCTTCGACAGTGACCTGA	584	
Db	942	AACTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCATCTTCGACAGTGACCTGA	1001	
Qy	585	TGGATAGTCCAGGCTGGGGCGCCCTTGCCCCCCCCAGAGTAACTCGGGCATCAGCGCCA	644	
Db	1002	TGGATAGTCCAGGCTGGGGCGCCCTTGCCCCCCCCAGAGTAACTCGGGCATCAGCGCCA	1061	
Qy	645	CGTGCTACGGCAGCGCGGGCGCATGGAGGGCCCGCCACCTACAGCGAGGTCATCG	704	
Db	1062	CGTGCTACGGCAGCGCGGGCGCATGGAGGGCCCGCCACCTACAGCGAGGTCATCG	1121	
Qy	705	GCCACTACCCGGGTCTCTTCCAGCACCATCGAGAGAGTGGGCCCTCCTTGTCTGG	764	
Db	1122	GCCACTACCCGGGTCTCTTCCAGCACCATCGAGAGAGTGGGCCCTCCTTGTCTGG	1181	
Qy	765	AGGGACCCGGTCCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA	824	
Db	1182	AGGGACCCGGTCCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA	1241	
Qy	825	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGCTGGG	884	
Db	1242	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGCTGGG	1301	
Qy	885	CTGCGTAGGTGAAAAGGCAG	904	
Db	1302	CTGCGTAGGTGAAAAGGCAG	1321	

RESULT 6
AX392419
LOCUS AX392419 861 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 3 from Patent WO0216416.

ACCESSION AX392419
VERSION AX392419.1 GI:19700734
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and
AUTHORS Turi,T.G.
TITLE Diagnosis and treatment of cardiovascular conditions
JOURNAL Patent: WO 0216416-A 3 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES Location/Qualifiers
source
1..861
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..>861
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD29006.1"
/db_xref="GI:19700735"
/translation="MHRLMGVNSTAAAAAGQPNVSTCNCKRSLFQSMETLEFVQI
IIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSESTVSGN
GIPEPQVYAPRPTDLAVPPFAQRERFHFQTPYLQHEIDLPTTISLSDGEEPPP
YQGPCTQLRDPHQLELNRESVRAPNRTIFDSLMDLSARLGGPCPPSSNSGISATC
YSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWS
KEKDKQKGHPL"
BASE COUNT 172 a 307 c 253 g 129 t
ORIGIN
Query Match 66.1%; Score 701; DB 6; Length 861;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 751; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 105 CGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGGTGATGATGATGGTGGTGG 164
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
110 CGGAGCTGGAGTTGTTTCAGATCATCATCGTGGTGGTGATGATGATGGTGGTGG 169
QY 165 TGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCACA 224
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
170 TGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCACA 229
QY 225 GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGCCCTCGGAGA 284
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
230 GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGCCCTCGGAGA 289
QY 285 GCACAGTGTAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCG 344
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
290 GCACAGTGTAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCG 349
QY 345 ACCGCCTGGCGTGGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACT 404
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
350 ACCGCCTGGCGTGGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACT 409
QY 405 ATCCGTACCTGCAGCACGAGATCGACCTGCGCCGCCCCACCACCATCTCGTGTGACACGGGAGG 464
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
410 ATCCGTACCTGCAGCACGAGATCGACCTGCGCCGCCCCACCACCATCTCGTGTGACACGGGAGG 469
QY 465 AGCCCCCACTACAGGGCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
470 AGCCCCCACTACAGGGCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 529
QY 525 AACTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGA 584
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
530 AACTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGA 589
QY 585 TGGATAGTGCAGGCTGGGCGGCCCTTGCCCCCCCCCAGCAGTAACTCGGGCGATCAGCGCCA 644
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
590 TGGATAGTGCAGGCTGGGCGGCCCTTGCCCCCCCCCAGCAGTAACTCGGGCGATCAGCGCCA 649
QY 645 CGTGCTACGCGAGCGGGCGGCGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTATCG 704

Db 650 CGTGCTACGGCAGCGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTATCG 709
QY 705 GCCACTACCCGGGTCTCTCCAGCACACGAGCAGTGGGCGCCCTCTCTGCTGG 764
Db 710 GCCACTACCCGGGTCTCTCTCCAGCACACGAGCAGTGGGCGCCCTCTCTGCTGG 769
QY 765 AGGGGACCCGGTCCACCACACACATCGCGGCCCTAGAGAGCGCAGCCATCTGGAGCA 824
Db 770 AGGGGACCCGGTCCACCACACACATCGCGGCCCTAGAGAGCGCAGCCATCTGGAGCA 829
QY 825 AAGAGAAGGATAAACAGAAAGGACACCCCTCTC 856
Db 830 AAGAGAAGGATAAACAGAAAGGACACCCCTCTC 861
RESULT 7
AF305426 61505 bp DNA linear PRI 12-OCT-2001
LOCUS Homo sapiens solid tumor-associated 1 protein (STAG1/PMEPA1) gene,
DEFINITION complete cds.
ACCESSION AF305426
VERSION AF305426.1 GI:15824468
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 61505)
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
Identification and characterization of a novel gene, STAG1,
up-regulated in renal cell carcinoma and other solid tumours
Unpublished
2 (bases 1 to 61505)
Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
Direct Submission
Submitted (13-SEP-2000) Centre for Molecular Biotechnology,
Queensland University of Technology, 2 George St, Brisbane, QLD
4001, Australia
FEATURES Location/Qualifiers
source
1..61505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2-q13.33"
1..61505
/gene="STAG1/PMEPA1"
join(1..429,50206..50360,56817..56870,57305..61505)
/gene="STAG1/PMEPA1"
/product="solid tumor-associated 1 protein"
join(321..429,50206..50360,56817..56870,57305..57850)
/gene="STAG1/PMEPA1"
/codon_start=1
/product="solid tumor-associated 1 protein"
/protein_id="AAL09357.1"
/db_xref="GI:15824469"
/translation="MHRLMGVNSTAAAAAGQPNVSTCNCKRSLFQSMETLEFVQI
IIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSGCLWPSESTVSGN
GIPEPQVYAPRPTDLAVPPFAQRERFHFQTPYLQHEIDLPTTISLSDGEEPPP
YQGPCTQLRDPHQLELNRESVRAPNRTIFDSLMDLSARLGGPCPPSSNSGISATC
YSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWS
KEKDKQKGHPL"
BASE COUNT 13329 a 16074 c 17430 g 14672 t
ORIGIN
Query Match 65.5%; Score 695; DB 9; Length 61505;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 312 AGCGCAGGTCTACGCCCCCGCTCGGCCACCGACCGCTGGCGCTCGCCCTTCGCC 371
Db 57303 AGCGCAGGTCTACGCCCCCGCTCGGCCACCGACCGCTGGCGCTCGCCCTTCGCC 57362

repeat_region 3378. .3426 /note="L1MB4 repeat: matches 6088. .6136 of consensus"
repeat_region 3799. .3896 /note="Charliel repeat: matches 681. .781 of consensus"
repeat_region 5331. .5793 /note="MLT1B repeat: matches 14. .466 of consensus"
repeat_region 5797. .5988 /note="6 copies 32 mer 86% conserved"
repeat_region 5855. .5978 /note="31 copies 4 mer gcac 61% conserved"
repeat_region 6474. .6591 /note="MIR repeat: matches 91. .218 of consensus"
repeat_region 6592. .6723 /note="FLAM_C repeat: matches 1. .132 of consensus"
repeat_region 6952. .7021 /note="L2 repeat: matches 2637. .2705 of consensus"
repeat_region 7358. .7671 /note="AluSx repeat: matches 1. .312 of consensus"
repeat_region 8521. .8554 /note="Alu repeat: matches 1. .34 of consensus"
repeat_region 8863. .9217 /note="MLT1D repeat: matches 105. .505 of consensus"
repeat_region 9228. .9746 /note="L2 repeat: matches 1459. .2028 of consensus"
repeat_region 9880. .10089 /note="L2 repeat: matches 2292. .2511 of consensus"
repeat_region 10094. .10206 /note="AluJb repeat: matches 188. .300 of consensus"
repeat_region 10263. .10573 /note="AluJb repeat: matches 1. .308 of consensus"
repeat_region 10658. .10753 /note="L2 repeat: matches 2647. .2749 of consensus"
repeat_region 11816. .12380 /note="LTR19B repeat: matches 1. .580 of consensus"
repeat_region 12519. .12813 /note="L1MD2 repeat: matches 6032. .6331 of consensus"
repeat_region 13559. .13726 /note="3 copies 56 mer 73% conserved"
repeat_region 13585. .13716 /note="22 copies 6 mer ctctct 67% conserved"
repeat_region 13586. .13717 /note="66 copies 2 mer tc 68% conserved"
repeat_region 13588. .13715 /note="4 copies 32 mer 78% conserved"
repeat_region 13589. .13724 /note="34 copies 4 mer cttt 77% conserved"
misc_feature complement(13622. .14142) /note="match: GSS: Em:AQ592603"
repeat_region 14313. .14437 /note="LTR16C repeat: matches 257. .386 of consensus"
repeat_region 15015. .15110 /note="3 copies 32 mer 79% conserved"
repeat_region 15021. .15110 /note="45 copies 2 mer ac 80% conserved"
repeat_region 15026. .15109 /note="7 copies 12 mer 82% conserved"
repeat_region 15028. .15111 /note="14 copies 6 mer cacaca 82% conserved"
repeat_region 15029. .15108 /note="20 copies 4 mer acac 83% conserved"
repeat_region 15273. .15399 /note="L2 repeat: matches 2159. .2285 of consensus"
repeat_region 15420. .15721 /note="AluSx repeat: matches 1. .303 of consensus"
repeat_region 16333. .16396 /note="2 copies 32 mer 98% conserved"
repeat_region 16575. .16690 /note="L2 repeat: matches 2596. .2711 of consensus"
repeat_region 18049. .18169 /note="MIR repeat: matches 86. .211 of consensus"
repeat_region 18312. .18438 /note="MIR repeat: matches 138. .250 of consensus"
repeat_region 18585. .18776

repeat_region /note="MER20 repeat: matches 7. .217 of consensus"
repeat_region 19211. .19290 /note="MIR repeat: matches 72. .154 of consensus"
repeat_region 20413. .20463 /note="MIR repeat: matches 90. .143 of consensus"
misc_feature complement(21005. .21672) /note="match: GSS: Em:AQ748384"
misc_feature complement(21187. .21672) /note="match: GSS: Em:AQ776209"
misc_feature complement(21278. .21663) /note="match: GSS: Em:AQ136459"
misc_feature complement(21192. .21672) /note="match: GSS: Em:AQ892114"
repeat_region 21667. .21775 /note="MIR repeat: matches 73. .192 of consensus"
repeat_region 23577. .23646 /note="2 copies 35 mer 100% conserved"
repeat_region 24257. .24591 /note="L2 repeat: matches 2336. .2710 of consensus"
repeat_region 25580. .25611 /note="16 copies 2 mer tc 87% conserved"
repeat_region 26333. .26643 /note="AluY repeat: matches 1. .310 of consensus"
repeat_region 27603. .27684 /note="L2 repeat: matches 2661. .2739 of consensus"
repeat_region 27823. .28041 /note="MIR repeat: matches 38. .242 of consensus"
repeat_region 28129. .28259 /note="FLAM_C repeat: matches 1. .127 of consensus"
repeat_region 31203. .31356 /note="L1MB5 repeat: matches 6015. .6176 of consensus"
repeat_region 32224. .32522 /note="AluSx repeat: matches 1. .298 of consensus"

Query Match 65.5%; Score 695; DB 9; Length 130435;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 745; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 312 AGCGCAGGTCTACGCCCCGCTCGGCCCCACCGACCGCCCTGCGCCCTTCGCCC 371
Db 128443 AGCGCAGGTCTACGCCCCGCTCGGCCCCACCGACCGCCCTGCGCCCTTCGCCC 128384

Qy 372 AGCGGAGCGCTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACC 431
Db 128383 AGCGGAGCGCTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACC 128324

Qy 432 TGCCGCCCCACCATCTCGCTGTCTAGACGGGAGGAGCCCCACCCCTACCCAGGGCCCCCTGCA 491
Db 128323 TGCCACCCACCATCTCGCTGTCTAGACGGGAGGAGCCCCACCCCTACCCAGGGCCCCCTGCA 128264

Qy 492 CCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGCGCGCAC 551
Db 128263 CCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGCGCGCAC 128204

Qy 552 CCCCCCAGAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGCTGGCGGCCCT 611
Db 128203 CCCCCCAGAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGCTGGCGGCCCT 128144

Qy 612 GCCCCCAGCAGTAACCTCGGCATCAGCGCCACGTCGTACGGCAGCGCGGCGCATGG 671
Db 128143 GCCCCCAGCAGTAACCTCGGCATCAGCGCCACGTCGTACGGCAGCGCGGCGCATGG 128084

Qy 672 AGGGCCGCGCCCCACCTACAGGAGGTCTATCGGCCACTACCCGGGGTCTCTCTCCAGC 731
Db 128083 AGGGCCGCGCCCCACCTACAGGAGGTCTATCGGCCACTACCCGGGGTCTCTCTCCAGC 128024

Qy 732 ACCAGCAGCAGTGGGCGCCCTCTTGTGTGGAGGGAGCCCGGCTCCACCACACACA 791
Db 128023 ACCAGCAGCAGTGGGCGCCCTCTTGTGTGGAGGGAGCCCGGCTCCACCACACACA 127964

Qy 792 TCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAAAGGACACC 851
Db 127963 TCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAAAGGACACC 127904

QY 852 CTCTCTAGGGTCCCCAGGGGGCGGGGCTGGGGCTGGAGTGGTGAAGGAGCAACTC 911
|||||
Db 127903 CTCTCTAGGGTCCCCAGGGGGCGGGGCTGGGGCTGGAGTGAAGGAGCAACTC 127844

QY 912 CGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCC 971
|||||
Db 127843 CGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCC 127784

QY 972 CTCCTCTCCACCTCCCTGTGTATATAATTTACATGTGATGTCTGGTCTGAATGCACAA 1031
|||||
Db 127783 CTCCTCTCCACCTCCCTGTGTATATAATTTACATGTGATGTCTGGTCTGAATGCACAA 127724

QY 1032 GCTAAGAGAGCTTGCAAAAAA 1057
|||||
Db 127723 GCTAAGAGAGCTTGCAAAAAA 127698
|||||

RESULT 9
AX593655
LOCUS AX593655 1583 bp DNA linear PAT 13-FEB-2003
DEFINITION Sequence 32 from Patent WO0242776.
ACCESSION AX593655
VERSION AX593655.1 GI:28375034
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Sun,Y., Recipon,H., Chen,S.Y. and Liu,C.
TITLE Compositions and methods relating to prostate specific genes and proteins
JOURNAL Patent: WO 0242776-A 32 30-MAY-2002;
Diadexus, Inc. (US)
FEATURES
source Location/Qualifiers
1 .1583
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 361 a 469 c 457 g 296 t
ORIGIN

Query Match 48.0%; Score 509; DB 6; Length 1583;
Best Local Similarity 99.8%; Pred. No. 7.2e-255;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 313 GCGCAGGTCTACGCCCCCGCTCGGCCCCACCGACCGCCTGGCCGTGCCGCCCTTCGCCCA 372
|||||
Db 827 GCGCAGGTCTACGCCCCCGCTCGGCCCCACCGACCGCCTGGCCGTGCCGCCCTTCGCCCA 886

QY 373 GCGGAGCGCTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCT 432
|||||
Db 887 GCGGAGCGCTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCT 946

QY 433 GCGGCCACCATCTCGTGTGACAGCGGGAGAGCCCCACCCCTACCAGGGCCCTGCAC 492
|||||
Db 947 GCGGCCACCATCTCGTGTGACAGCGGGAGAGCCCCACCCCTACCAGGGCCCTGCAC 1006

QY 493 CTTCCAGCTTCGGGACCCCGAGCAGAGCTGGAACCTGAACCGGGAGTCGGTGCAGCACC 552
|||||
Db 1007 CTTCCAGCTTCGGGACCCCGAGCAGAGCTGGAACCTGAACCGGGAGTCGGTGCAGCACC 1066

QY 553 CCAAAACAGAACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGCGGCCCTTG 612
|||||
Db 1067 CCAAAACAGAACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGCGGCCCTTG 1126

QY 613 CCCCCCAGCAGTAACCTCGGCATCAGCGCCACGTGTACGGCAGCGCGCGGCATGGA 672
|||||
Db 1127 CCCCCCAGCAGTAACCTCGGCATCAGCGCCACGTGTACGGCAGCGCGCGGCATGGA 1186

QY 673 GGGGCGCGCCACCTACAGCAGGTTCATCGGCCACTACCGGGGTCTCTCCAGCA 732
|||||

Db 1187 GGGCGCGCGCCACCTACAGCGAGGTTCATCGGCCACTACCGGGGTCTCTTCCAGCA 1246

QY 733 CCAGCAGAGCAGTGGCGCCCTCTCTTGTCTGGAGGGGACCCGGCTCCACCACACACAT 792
|||||
Db 1247 CCAGCAGAGCAGTGGCGCCCTCTCTTGTCTGGAGGGGACCCGGCTCCACCACACACAT 1306

QY 793 CGCGCCCTTAGAGAGCGCAGCCCATCTGGAGCAAAGAGAAGGATAAACAGAAAGGACCC 852
|||||
Db 1307 CGCGCCCTTAGAGAGCGCAGCCCATCTGGAGCAAAGAGAAGGATAAACAGAAAGGACCC 1366

QY 853 TCTCTAGGTCCCCAGGGG 872
|||||
Db 1367 TCTCTAGGTCCCCAGGGG 1386
|||||

RESULT 10
AX071267
LOCUS AX071267 408 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 1739 from Patent WO0102568.
ACCESSION AX071267
VERSION AX071267.1 GI:12581618
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J., Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D., Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M., Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1739 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1 .408
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 81 a 159 c 115 g 53 t
ORIGIN

Query Match 33.0%; Score 350; DB 6; Length 408;
Best Local Similarity 99.8%; Pred. No. 1.5e-171;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 166 GATCACGTGCCTGTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGCACAG 225
|||||
Db 8 GATCACGTGCCTGTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGCACAG 67

QY 226 CCAGGGCGGAGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGCCCTCGGAGAG 285
|||||
Db 68 CCAGGGCGGAGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGCCCTCGGAGAG 127

QY 286 CACAGTGTCAAGCAACGGAATCCCAGAGCCGCGAGGTCTACGCCCCGCCCTCGGCCACCGA 345
|||||
Db 128 CACAGTGTCAAGCAACGGAATCCCAGAGCCGCGAGGTCTACGCCCCGCCCTCGGCCACCGA 187

QY 346 CCGCCTGGCCGTGCCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCACTA 405
|||||
Db 188 CCGCCTGGCCGTGCCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCCACTA 247

QY 406 TCCGTACCTGCAGCAGAGATCGACCTGCCGCCACCATCTCGCTGTGACACGGGAGGA 465
|||||
Db 248 TCCGTACCTGCAGCAGAGATCGACCTGCCACCACCATCTCGCTGTGACACGGGAGGA 307

QY 466 GCCCCCACTACAGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGA 525
|||||
Db 308 GCCCCCACTACAGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGA 367

QY 526 ACTGAACCGGGAGTCGGTGGCGGCACCCCCCAACAGAACCA 566
|||||
Db 368 ACTGAACCGGGAGTCGGTGGCGGCACCCCCCAACAGAACCA 408
|||||

RESULT 11
AX011709
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .812
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 157 a 272 c 237 g 146 t
ORIGIN
Query Match 30.3%; Score 322; DB 6; Length 812;
Best Local Similarity 100.0%; Pred. No. 7e-157;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 168 TCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCC 227
|||||
Db 426 TCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCC 485
QY 228 AGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAGGATGCTGTGGCCCTCGGAGAGCA 287
|||||
Db 486 AGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAGGATGCTGTGGCCCTCGGAGAGCA 545
QY 288 CAGTGTCAAGCAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCCTCGGCCACCGACC 347
|||||
Db 546 CAGTGTCAAGCAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCCTCGGCCACCGACC 605
QY 348 GCCTGGCCGTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACGGCTTCCAGCCCACTATC 407
|||||
Db 606 GCCTGGCCGTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACGGCTTCCAGCCCACTATC 665
QY 408 CGTACTGCAGCAGCAGATCGACCTGCCGCCCCACCATCTCGTGTGTCAGACGGGAGGAGC 467
|||||
Db 666 CGTACTGCAGCAGCAGATCGACCTGCCGCCCCACCATCTCGTGTGTCAGACGGGAGGAGC 725
QY 468 CCCACCCCTACGAGGCCCTG 489
|||||
Db 726 CCCACCCCTACGAGGCCCTG 747

RESULT 12
AX392430/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
AX392430
Sequence 14 from Patent WO0216416.
AX392430
AX392430.1 GI:19700746
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and
Turi,T.G.
Diagnosis and treatment of cardiovascular conditions

JOURNAL Patent: WO 0216416-A 14 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source
Location/Qualifiers
1. .693
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 101 a 205 c 237 g 149 t 1 others
ORIGIN
Query Match 29.5%; Score 313; DB 6; Length 693;
Best Local Similarity 99.3%; Pred. No. 3.6e-152;
Matches 583; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 468 CCCACCCCTACGAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC 527
|||||
Db 587 CCCACCCCTACGAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC 528
QY 528 TGAACCGGAGTCGGTGGCGCACCCCAACAGAACATCTTCGACAGTGACCTGATGG 587
|||||
Db 527 TGAACCGGAGTCGGTGGCGCACCCCAACAGAACATCTTCGACAGTGACCTGATGG 468
QY 588 ATAGTGCCAGGCTGGGGCCCTGCCCGCCCGCAGAGTAACCTCGGCGCATCGGCCACGT 647
|||||
Db 467 ATAGTGCCAGGCTGGGGCCCTGCCCGCCCGCAGAGTAACCTCGGCGCATCGGCCACGT 408
QY 648 GCTACGGCAGCGCGCGGCATGGAGGGCGCGCCCGCCACCTACAGCGAGGTCTATCGGCC 707
|||||
Db 407 GCTACGGCAGCGCGCGGCATGGAGGGCGCGCCCGCCACCTACAGCGAGGTCTATCGGCC 348
QY 708 ACTACCGGGGTCTCTTCCAGCACACAGCAGAGAGTGGGGCGCCCTCTTGTGGA-G 766
|||||
Db 347 ACTACCGGGGTCTCTTCCAGCACACAGCAGAGAGTGGGGCGCCCTCTTGTGGAAG 288
QY 767 GGGACCGGGTCCACCAACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAA 826
Db 287 GGGACCGGGTTCGCCCAACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAA 228
QY 827 GAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCGGGCTGGGGCT 886
|||||
Db 227 GAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCGGGCTGGGGCT 168
QY 887 GCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGG 946
|||||
Db 167 GCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGG 108
QY 947 GGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACA 1006
|||||
Db 107 GGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACA 48
QY 1007 TGTGATGTCTGGTCTGAATGCACAAAGCTAAGAGAGCTTGCAAAAAA 1053
|||||
Db 47 TGTGATGTCTGGTCTGAATGCACAAAGCTAAGAGAGCTTGCAAAAAA 1

RESULT 13
HSJ1059L7/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
HSJ1059L7
Human DNA sequence from clone RP5-1059L7 on chromosome
20q13.2-13.33 Contains the 5' end of the TMEPAI (PMEPAI) gene
encoding an androgen induced 1b transmembrane protein, ESTs, STSs,
GSSs and two CpG islands, complete sequence.
AL121913
AL121913.4 GI:7161781
HTG; CpG island; PMEPAI; TMEPAI; transmembrane protein.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 150224)
Skuce,C.
Direct Submission
Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Mar 6, 2000 this sequence version replaced gi:7007305. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP5-1059L7 The true left end of clone RP11-402F1 is at 106677 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-1059L7 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES

source

Location/Qualifiers

1. .150224

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="20"

/map="q13.2-13.33"

/clone="RP5-1059L7"

/clone_lib="RPCI-5"

283. .355

/note="MIR repeat: matches 71. .145 of consensus"

1509. .1628

/note="MIR repeat: matches 46. .165 of consensus"

2076. .2155

/note="L2 repeat: matches 2616. .2707 of consensus"

complement(2445. .53425)

/gene="TMEPAI"

complement(join(<2445. .2599,53362. .53425))

/gene="TMEPAI"

/product="dJ1059L7.1.2 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 2)"

transmembrane protein (PMEPA1), isoform 2)"

match: ESTs: Em:AA088767"

/evidence=not_experimental

complement(join(<2445. .2599,53362. .53365))

/gene="TMEPAI"

/note="continues in dJ718J7 (AL035541)"

/codon_start=1

/evidence=not_experimental

/product="dJ1059L7.1.2 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 2)"

/protein_id="CAC32857.1"

/db_xref="GI:13160408"

/translation="MABLEFVQIIIIIVVMVMVVITCLLSHYKLSARSPISRHSQG

RRREDALSS"

complement(join(<2445. .2599,52376. .52477))

/gene="TMEPAI"

/product="dJ1059L7.1.1 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 1)"

/note="match: cDNAs: Em:AF009426 Em:AF009427 Em:AF009425

Em:AF009424

match: ESTs: Em:AA249792 Em:AI594390 Em:AA128075

CDS

Em:AA088767"

/evidence=not_experimental

complement(join(<2445. .2599,52376. .52472))

/gene="TMEPAI"

/note="continues in dJ718J7 (AL035541)

match: proteins: Tr:O15166 Tr:O15168"

/codon_start=1

/evidence=not_experimental

/product="dJ1059L7.1.1 (androgen induced type 1b

transmembrane protein (PMEPA1), isoform 1)"

/protein_id="CAB88144.1"

/db_xref="GI:7619746"

/db_xref="SPTREMBL:Q9NTR9"

/translation="MGVNSTAAAAAGQPNVSVCTCNCKRSLFQSMETELFVQIIIIIV

VMMVMVVITCLLSHYKLSARSPISRHSQRRREDALSS"

2538. .2567

/note="10 copies 3 mer cac 90% conserved"

3128. .3606

/note="match: GSS: Em:AQ059507"

3214. .3324

/note="3 copies 37 mer 83% conserved"

3339. .3413

/note="3 copies 25 mer 92% conserved"

3390. .3921

/note="19 copies 28 mer 61% conserved"

3394. .3893

/note="25 copies 20 mer 59% conserved"

3407. .3892

/note="27 copies 18 mer 58% conserved"

3421. .3910

/note="7 copies 70 mer 62% conserved"

3473. .3904

/note="12 copies 36 mer 58% conserved"

3553. .3932

/note="10 copies 38 mer 64% conserved"

3573. .3697

/note="5 copies 25 mer 76% conserved"

3746. .3921

/note="8 copies 22 mer 64% conserved"

4032. .4069

/note="19 copies 2 mer tt 81% conserved"

4792. .5001

/note="3 copies 70 mer 82% conserved"

5002. .5190

/note="MER20 repeat: matches 1. .200 of consensus"

5224. .5282

/note="MIR repeat: matches 94. .154 of consensus"

8103. .8205

/note="MIR repeat: matches 48. .161 of consensus"

complement(8858. .9277)

/gene="TMEPAI"

/note="match: GSS: Em:AQ881699"

9224. .9382

/note="MIR repeat: matches 81. .239 of consensus"

9286. .9706

/note="match: GSS: Em:AQ565377"

9295. .9861

/note="match: GSS: Em:AQ757992"

10035. .10140

/note="Charlie4 repeat: matches 39. .146 of consensus"

11175. .11343

/note="MIR repeat: matches 11. .196 of consensus"

11630. .12037

/note="L2 repeat: matches 2290. .2705 of consensus"

12054. .12362

/note="AluX repeat: matches 1. .311 of consensus"

14448. .14493

/note="23 copies 2 mer at 76% conserved"

14634. .14813

/note="MER20 repeat: matches 1. .187 of consensus"

15096. .15497

/note="CpG island"

/evidence=not_experimental


```

repeat_region 15690. .15935
/note="MIR repeat: matches 7. .254 of consensus"
repeat_region 16989. .17387
/note="MLT1A2 repeat: matches 1. .374 of consensus"
repeat_region 17647. .17703
/note="L2 repeat: matches 2679. .2736 of consensus"
misc_feature 17699. .18179
/note="match: GSS: Em:AQ703107"
repeat_region 17902. .18039
/note="MIR repeat: matches 107. .243 of consensus"
repeat_region 19070. .19291
/note="MIR repeat: matches 35. .261 of consensus"
repeat_region 19954. .20005
/note="L2 repeat: matches 2358. .2411 of consensus"
repeat_region 20264. .20432
/note="L2 repeat: matches 2581. .2750 of consensus"
repeat_region 20530. .21282
/note="L1MD2 repeat: matches 5595. .6341 of consensus"
repeat_region 21980. .22192
/note="L2 repeat: matches 2322. .2562 of consensus"
repeat_region 23102. .23203
/note="MIR repeat: matches 117. .218 of consensus"
repeat_region 23231. .23438
/note="MER20 repeat: matches 3. .218 of consensus"
repeat_region 23439. .23512
/note="MIR repeat: matches 182. .257 of consensus"
repeat_region 24460. .24580
/note="L1MD1 repeat: matches 6102. .6221 of consensus"
repeat_region 24752. .25042
/note="L2 repeat: matches 2386. .2709 of consensus"
repeat_region 25208. .25295
/note="MIR repeat: matches 61. .155 of consensus"
repeat_region 26450. .26485
/note="18 copies 2 mer tg 97% conserved"
repeat_region 29374. .29501
/note="MIR repeat: matches 55. .190 of consensus"
repeat_region 29550. .29728
/note="MIR repeat: matches 69. .248 of consensus"

```

	Query Match	14.8%;	Score 157;	DB 9;	Length 150224;
	Beat Local Similarity	100.0%;	Pred. No. 1.6e-70;		
	Matches 157;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	104	GCGGAGCTGGAGTTGTTCAGATCATCATCGTGGTGATGATGGTGGTG	163		
Dd	2600	GCGGAGCTGGAGTTGTTCAGATCATCATCGTGGTGATGATGGTGGTG	2541		
QY	164	GTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCTTCATCAGCCGGCAC	223		
Dd	2540	GTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCTTCATCAGCCGGCAC	2481		
QY	224	AGCCAGGGCGGAGAGAGAAGATGCCCTGTCCTCAG	260		
Dd	2480	AGCCAGGGCGGAGAGAGAAGATGCCCTGTCCTCAG	2444		

RESULT 14
AK056098/c

LOCUS	AK056098	2570 bp	mRNA	linear	PRI 01-AUG-2002
DEFINITION	Homo sapiens cDNA FLJ31536 fis, clone NT2RI2000689.				
ACCESSION	AK056098				
VERSION	AK056098.1 GI:16551411				
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota;	Chordata;	Cranialia;	Vertebrata;	Euteleostomi;
	Mammalia;	Eutheria;	Primates;	Catarrhini;	Hominidae; Homo.
REFERENCE	1				
AUTHORS	Ninomiya,K.,	Wagatsuma,M.,	Kanda,K.,	Kondo,H.,	Yokoi,T.,
	Kodaira,H.,	Furiya,T.,	Takahashi,M.,	Kikkawa,E.,	Omura,Y., Abe,K.,
	Kamihara,K.,	Katsuta,N.,	Sato,K.,	Tanikawa,M.,	Yamazaki,M.,
	Sugiyama,T.,	Irie,R.,	Otsuki,T.,	Sato,H.,	Wakamatsu,A., Ishii,S.,
	Yamamoto,J.,	Isono,Y.,	Kawai-Hio,Y.,	Saito,K.,	Nishikawa,T.,

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
NEDO human cDNA sequencing project	Unpublished	2 (bases 1 to 2570)	Isogai, T., Otsuki, T. and Sugiyama, T.	Submitted (24-OCT-2001)	Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology (RAB); CDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES	source
Location/Qualifiers	
1. .2570	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="NT2RI2000689"	
/cell_line="NT2"	
/cell_type="teratocarcinoma"	
/clone_lib="NT2RI2"	
/note="cloning vector: pME18SFL3~mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"	
BASE COUNT	611 a 675 c 712 g 572 t
ORIGIN	

```

QY      61  TGGAGAACTGAAGGCGGAC  79
      |||||||||||||||
Db      321  TGGAGAACTGAAGGCGGAC  303

RESULT 15
AX199565
LOCUS      AX199565      51 bp      DNA      linear      PAT 29-AUG-2001
DEFINITION      Sequence 495 from Patent WO0151670.
ACCESSION      AX199565
VERSION      AX199565.1  GI:15389996
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Shimkets, R.A. and Leach, M.D.
TITLE      Nucleic acids containing single nucleotide polymorphisms and
              methods of use thereof
JOURNAL      Patent: WO 0151670-A 495 19-JUL-2001;
              Curagen Corporation (US)
FEATURES
              Location/Qualifiers
                1..51
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
              misc_feature
                /note="1 of 2 allelic variants (496 is other entry)"
                Accession number cg42747251"

```


BASE COUNT 12 a 16 c 18 g 5 t
ORIGIN

Query Match 4.8%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCC 251
|||
Db 1 CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAGATGCC 51
|||

Search completed: December 8, 2003, 05:33:47
Job time : 4149 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 02:02:00 ; Search time 344 Seconds
(without alignments)
8325.883 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tctctctgggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :				N_Geneseq 19Jun03:*	
				1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*	
				2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*	
				3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*	
				4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*	
				5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*	
				6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*	
				7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*	
				8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*	
				9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*	
				10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*	
				11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*	
				12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*	
				13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*	
				14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*	
				15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*	
				16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*	
				17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*	
				18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*	
				19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*	
				20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*	
				21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*	
				22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*	
				23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*	
				24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*	
				25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1061	100.0	1061	21 AAA47429	Sequence encoding
2	1056	99.5	1334	25 ABZ36103	Human secretory po
3	961	90.6	969	21 AAA75151	CDNA encoding a hu
4	952	89.7	969	21 AAA75163	CDNA clone encodin
5	926	87.3	1140	24 ABK92120	Prostate cancer-as
6	910	85.8	969	21 AAA75164	CDNA clone encodin
7	910	85.8	969	21 AAA75165	CDNA clone encodin
8	891	84.0	1066	22 AAI57868	Human polynucleoti

9	749	70.6	1321	24	ABK12137	Human CDNA encodin
10	548	51.6	1069	22	AAI59654	Human polynucleoti
11	509	48.0	1583	24	ABS61424	Prostate specific
12	415	39.1	474	25	ABZ84732	Toxicologically re
13	350	33.0	408	22	AAF65983	Novel human polynu
14	322	30.3	812	20	AAZ52964	Human prostate tum
15	313	29.5	693	24	ABK12143	Human MIVR-1 homol
16	229	21.6	254	21	AAA41265	Human secreted exp
17	93	8.8	426	23	AAS84502	DNA encoding novel
18	91	8.6	1879	23	AAS84503	DNA encoding novel
19	68	6.4	522	24	ABT10027	Human breast cance
20	60	5.7	60	24	ABN40872	Human spliced tran
21	51	4.8	51	22	AAH89714	Human coding sequ
22	44	4.1	878	24	ABK12142	Mouse CDNA encodin
23	44	4.1	1713	21	AAA75152	CDNA encoding a mu
24	44	4.1	1713	21	AAA75166	CDNA clone encodin
25	44	4.1	1713	21	AAA75167	CDNA clone encodin
26	44	4.1	1713	21	AAA75168	CDNA clone encodin
27	42	4.0	673	24	ABT09178	Phase-1 Rat CT gen
28	32	3.0	577	22	ABA50270	Human breast cell
29	32	3.0	577	22	ABA68210	Human foetal liver
30	32	3.0	577	22	ABA35222	Probe #13688 for g
31	32	3.0	577	22	AAK16592	Human brain expres
32	32	3.0	577	22	AAK42346	Human bone marrow
33	32	3.0	577	22	AAI23118	Probe #13051 for g
34	32	3.0	577	22	AAI48425	Probe #17111 used
35	32	3.0	577	22	AAI08772	Probe #8763 used t
36	32	3.0	577	23	ABS41955	Human liver single
37	32	3.0	577	24	ABS16402	Human genome-deriv
38	32	3.0	1964	22	ABA45134	Human breast cell
39	32	3.0	1964	22	ABA55612	Human foetal liver
40	32	3.0	1964	22	ABA25310	Probe #3776 for ge
41	32	3.0	1964	22	AAK03841	Human brain expres
42	32	3.0	1964	22	AAK29307	Human bone marrow
43	32	3.0	1964	22	AAI13901	Probe #3834 for ge
44	32	3.0	1964	22	AAI35267	Probe #3953 used t
45	32	3.0	1964	22	AAI03769	Probe #3760 used t

ALIGNMENTS

RESULT 1

AAA47429

ID AAA47429 standard; DNA; 1061 BP.

XX

AC AAA47429;

XX

DT 20-OCT-2000 (first entry)

XX

DE Sequence encoding human neuron-associated protein.

XX

KW Neuron associated protein; NEUAP; neurological disorder; epilepsy;
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;
KW Alzheimer's disease; Pick's disease; Huntington's disease;
KW dementia; Parkinson's disease; demyelinating disease; meningitis;
KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;
KW cerebral palsy; muscular dystrophy; central nervous system; CNS;
KW peripheral nervous system; PNS; myopathy; schizophrenia;
KW actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;
KW AIDS; Addison's disease; adult respiratory distress syndrome;
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
KW Werner syndrome, trauma; human; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 101..859

/*tag= a

/product= Neuron associated protein

XX WO200034477-A2.
PN 15-JUN-2000.
XX 10-DEC-1999; 99WO-US30408.
XX 11-DEC-1998; 98US-0210083.
PR 11-DEC-1998; 98US-9123456.
PR 09-FEB-1999; 99US-0119365.
PR 16-MAR-1999; 99US-0124687.
XX (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Yue H, Baughn MR, Hillman JL, Lal P, Au-young J, Yang J;
PI Lu DAM, Azimzai Y;
XX
DR WPI; 2000-423423/36.
DR P-PSDB; AAB01388.
XX
PT New human neuron-associated proteins and polynucleotides encoding them,
PT useful for diagnosis, treatment and prevention of cell proliferative
PT disorders including cancer, neuronal and neurological disorders
XX
PS Claim 9; Page 136; 145pp; English.
XX
CC Human neuron-associated proteins (NEUP) can be used for for
CC treating or preventing a disorder associated with decreased
CC expression or activity of NEUP. Antagonists of NEUP are useful for
CC treating or preventing disorder associated with increased expression
CC or activity of NEUP. NEUP or their fragments or derivatives are
CC useful for treating neurological disorder such as epilepsy, ischemic
CC cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's
CC disease, Pick's disease, Huntington's disease, dementia and
CC Parkinson's disease. NEUPs are also useful for treating other
CC demyelinating diseases, bacterial and viral meningitis, prion
CC diseases including kuru, Creutzfeldt-Jakob disease, nutritional and
CC metabolic diseases of the nervous system, neurofibromatosis, other
CC developmental disorders of the central nervous system, cerebral
CC palsy, neuroskeletal disorders, autonomic nervous system disorders,
CC cranial nerve disorders, spinal cord diseases, muscular dystrophy and
CC other neuromuscular disorders, peripheral nervous system disorders,
CC inherited, metabolic, endocrine, and toxic myopathies, mental
CC disorders including mood, anxiety and schizophrenic disorders, a cell
CC proliferative disorder such as actinic keratosis, arteriosclerosis,
CC atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective
CC tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal
CC haemoglobinuria, cancers of the adrenal gland, bladder, bone,
CC bone marrow, breast, cervix, and an autoimmune/inflammatory
CC disorder such as acquired immunodeficiency syndrome (AIDS), Addison's
CC disease, adult respiratory distress syndrome, allergies, ankylosing
CC spondylitis, amyloidosis, anemia, asthma, Werner syndrome,
CC complications of cancer, hemodialysis, and extracorporeal circulation,
CC viral, bacterial, fungal parasitic, protozoal, and helminthic
CC infections, and trauma. This sequence was given the Incyte ID no.
XX 1871288CB1.
SQ Sequence 1061 BP; 225 A; 342 C; 326 G; 168 T; 0 other;

Query Match 100.0%; Score 1061; DB 21; Length 1061;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1061; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCTCCTTGGGTTGGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
DB 1 TCCTCCTTGGGTTGGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
QY 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAAACCAGGCAATGGCGGAGCTGGAGTTTGT 120
DB 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAAACCAGGCAATGGCGGAGCTGGAGTTTGT 120
QY 121 TCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180

Db 121 TCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
QY 181 GAGCCACTACAAGTGTCTGCACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAG 240
Db 181 GAGCCACTACAAGTGTCTGCACGGTCTTTCATCAGCCGGCACAGCCAGGGCGGAGGAG 240
QY 241 AGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAA 300
Db 241 AGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAA 300
QY 301 CGGAATCCCAGAGCCGCGAGGTCTACGCCCCCGCTCGGCCCCACCGCCGCTGGCCGTGCC 360
Db 301 CGGAATCCCAGAGCCGCGAGGTCTACGCCCCCGCTCGGCCCCACCGCCGCTGGCCGTGCC 360
QY 361 GCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCAGAGCCACCTATCCGTACCTGCAGCA 420
Db 361 GCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCAGAGCCACCTATCCGTACCTGCAGCA 420
QY 421 CGAGATCGACCTGCCGCCACCATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCA 480
Db 421 CGAGATCGACCTGCCGCCACCATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCA 480
QY 481 GGGCCCTGCACCCCTCGAGCTTCGGGACCCCGAGCAGCTGGAACCGGAGTC 540
Db 481 GGGCCCTGCACCCCTCGAGCTTCGGGACCCCGAGCAGCTGGAACCGGAGTC 540
QY 541 GGTGGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGGATAGTCCAGGCT 600
Db 541 GGTGGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGGATAGTCCAGGCT 600
QY 601 GGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGGCCACGTCTACGGCAGCGG 660
Db 601 GGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGGCCACGTCTACGGCAGCGG 660
QY 661 CGGGCGCATGGAGGGCGCCGCCACCTACAGCAGGTTCATCGGCCACTACCCGGGTC 720
Db 661 CGGGCGCATGGAGGGCGCCGCCACCTACAGCAGGTTCATCGGCCACTACCCGGGTC 720
QY 721 CTCCTTCAGCACCCAGCAGCAGTGGGGCGCCCTCTCTGCTGGAGGGACCCGGCTCCA 780
Db 721 CTCCTTCAGCACCCAGCAGCAGTGGGGCGCCCTCTCTGCTGGAGGGACCCGGCTCCA 780
QY 781 CCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA 840
Db 781 CCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA 840
QY 841 GAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGCGGCTGGGCTGGGTGAAAAG 900
Db 841 GAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGCGGCTGGGCTGGGTGAAAAG 900
QY 901 GCAGAACTACCCCGCTTCTTAGAAGAGGAGTGAAGAGCGGGGGGGCGCAGCAACGC 960
Db 901 GCAGAACTACCCCGCTTCTTAGAAGAGGAGTGAAGAGCGGGGGGGCGCAGCAACGC 960
QY 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTGTGTC 1020
Db 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTGTGTC 1020
QY 1021 TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1061
Db 1021 TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1061

RESULT 2
ABZ36103
ID ABZ36103 standard; cDNA; 1334 BP.
XX
AC ABZ36103;
XX
DT 10-FEB-2003 (first entry)
XX
DE Human secretory polynucleotide SPTM SEQ ID NO 267.
XX

Db 966 A 966

RESULT 4

AAA75163

ID AAA75163 standard; cDNA; 969 BP.

XX AAA75163;

AC AAA75163;

XX 15-JAN-2001 (first entry)

DT

XX cDNA clone encoding a human TANGO 261 polypeptide.

DE

XX

KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;

KW cellular proliferation; cellular differentiation; cellular adhesion;

KW von Willebrand factor-associated disorder; cell trafficking; cancer;

KW hematopoietic associated disease; atelectasis; pulmonary congestion;

KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;

KW intestinal disorder; spleen associated disease; renal disorder;

KW cardiovascular disorder; ischemic heart disease; hydrocephalus;

KW brain herniation; iatrogenic disease; inflammation; meningitis;

KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;

KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS 6..764

FT /*tag= a

FT /product= "TANGO 261"

XX

PN WO200052022-A1.

XX

PD 08-SEP-2000.

XX

PF 01-MAR-2000; 2000WO-US05226.

XX

PR 01-MAR-1999; 99US-0122458.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX

DR WPI; 2000-579269/54.

DR P-PSDB; AAB18461.

XX

PT Novel human and murine secreted proteins designated TANGO 216, 261,

PT 262, 266 and 267 useful as modulating agents of cellular processes,

PT e.g. for treating cancer -

XX

PS Disclosure; Page -; 175pp; English.

XX

CC AAA75163-65 encode human TANGO 261 proteins. The specification also

CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO

CC polypeptides can be used to modulate cellular proliferation, modulate

CC cellular differentiation and/or modulate cellular adhesion. The

CC proteins can be used to treat any von Willebrand factor-associated

CC disorder, regulate extracellular matrix structuring, cellular adhesion,

CC and cell trafficking and/or migration, modulate cellular interactions,

CC modulate cell adhesion in proliferative disorders, such as cancer,

CC modulate the proliferation, differentiation, and/or function of cells

CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood

CC and hematopoietic associated diseases and disorders, atelectasis,

CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial

CC asthma and bronchiectasis, intestinal disorders, spleen associated

CC diseases, modulate renal disorders, treat cardiovascular disorders such

CC as ischemic heart disease, modulate the proliferation, differentiation,

CC and/or function of bone and cartilage cells and to treat bone and/or

CC cartilage associated diseases or disorder. They may also be used to

CC treat disorders associated with the ovaries, and cerebral oedema,

CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,

CC bacterial and viral meningitis, Alzheimer's Disease, cerebral

CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,

CC hydrocephalus and encephalitis, and treat hepatic disorders.

CC note: the present sequence does not appear in the specification; it was

CC created using information provided.

XX

SQ Sequence 969 BP; 211 A; 317 C; 293 G; 148 T; 0 other;

Query Match 89.7%; Score 952; DB 21; Length 969;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 952; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 CTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGGTGGTGGTGATC 169

Db 15 CTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGGTGGTGGTGATC 74

QY 170 ACGTGCCTGTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCAG 229

Db 75 ACGTGCCTGTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCAG 134

QY 230 GGGCGGAGGAGAGATGCCCTGTCTCTCAGAAAGGATGCCCTGTGGCCCTCGGAGAGCACA 289

Db 135 GGGCGGAGGAGAGATGCCCTGTCTCTCAGAAAGGATGCCCTGTGGCCCTCGGAGAGCACA 194

QY 290 GTGTCAGGCAACGGAAATCCAGAGCCGCGAGGTCTACGCCCCCGCTCGGCCCCACCGACCGC 349

Db 195 GTGTCAGGCAACGGAAATCCAGAGCCGCGAGGTCTACGCCCCCGCTCGGCCCCACCGACCGC 254

QY 350 CTGGCCGTGCGGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCAGCCCACTATCCG 409

Db 255 CTGGCCGTGCGGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCAGCCCACTATCCG 314

QY 410 TACCTGCAGCACGAGATCGACCTGCCGCCCAACCATCTCGCTGTCTCAGACGGGAGGAGCC 469

Db 315 TACCTGCAGCACGAGATCGACCTGCCGCCCAACCATCTCGCTGTCTCAGACGGGAGGAGCC 374

QY 470 CCACCCCTACAGGGCCCTGCACTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAATG 529

Db 375 CCACCCCTACAGGGCCCTGCACTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAATG 434

QY 530 AACCGGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGAT 589

Db 435 AACCGGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGAT 494

QY 590 AGTGCCAGGCTGGGCGGCCCTGCCCCCCCAGCACTAACTCGGGCATCAGCGCCACGTGC 649

Db 495 AGTGCCAGGCTGGGCGGCCCTGCCCCCCCAGCACTAACTCGGGCATCAGCGCCACGTGC 554

QY 650 TACGGCAGCGGGCGGCATGGAGGGCGCCGCCCACTACAGCGAGGTCTATCGGCCAC 709

Db 555 TACGGCAGCGGGCGGCATGGAGGGCGCCGCCCACTACAGCGAGGTCTATCGGCCAC 614

QY 710 TACCGGGGTCTCTCTCAGCACACGAGCAGTGGGCGCCCTCTCTTGTGGAGGG 769

Db 615 TACCGGGGTCTCTCTCAGCACACGAGCAGTGGGCGCCCTCTCTTGTGGAGGG 674

QY 770 ACCCGGCTCCACCAACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAG 829

Db 675 ACCCGGCTCCACCAACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAG 734

QY 830 AAGGATAAACAGAAAGACACCTCTCTAGGGTCCCAGGGGGCGCGGTGGGCTGCG 889

Db 735 AAGGATAAACAGAAAGACACCTCTCTAGGGTCCCAGGGGGCGCGGTGGGCTGCG 794

QY 890 TAGGTGAAAAGGCAGAACTCCGCGCTTCTTAGAAGAGAGTGAAGGAAGCGGGGG 949

Db 795 TAGGTGAAAAGGCAGAACTCCGCGCTTCTTAGAAGAGAGTGAAGGAAGCGGGGG 854

QY 950 CGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATTTACATGT 1009

Db 855 CGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATTTACATGT 914

QY 1010 GATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1061

Db 915 GATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 966

RESULT 5
ABK92120
ID ABK92120 standard; DNA; 1140 BP.
XX
AC ABK92120;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated DNA sequence #6.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX gene therapy; gene; ds.
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US32045.
XX
PR 13-OCT-2000; 2000US-0687576.
PR 08-DEC-2000; 2000US-0733288.
PR 08-DEC-2000; 2000US-0733742.
PR 24-JAN-2001; 2001US-263957P.
PR 16-MAR-2001; 2001US-276791P.
PR 16-MAR-2001; 2001US-276888P.
PR 06-APR-2001; 2001US-281922P.
PR 24-APR-2001; 2001US-286214P.
PR 30-APR-2001; 2001US-0847046.
PR 04-MAY-2001; 2001US-288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WPI; 2002-471335/50.
DR P-PSDB; ABG61805.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue -
XX
PS Claim 22; Page 305; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate
CC cancer-associated transcript in a cell from a patient. The method
CC comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridise to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences.
XX
SQ Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;

Query Match 87.3%; Score 926; DB 24; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 81 GTCTCCTCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGG 140

Db 76 GTCTCCTCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGG 135
QY 141 TGGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 200
Db 136 TGGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 195
QY 201 CACGGTCTTCATCAGCGGCGCACAGCCAGGGCGGAGGAGAGAAAGATGCCCTGTCTCAG 260
Db 196 CACGGTCTTCATCAGCGGCGCACAGCCAGGGCGGAGGAGAGAAAGATGCCCTGTCTCAG 255
QY 261 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTACAGCAACGGAATCCAGAGCGGAGG 320
Db 256 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTACAGCAACGGAATCCAGAGCGGAGG 315
QY 321 TCTACGCCCCCGCTCGGCCACCGACCGCCTGGCCCTGGCCCTGGCCCTGGCCCTGGCCGAGC 380
Db 316 TCTACGCCCCCGCTCGGCCACCGACCGCCTGGCCCTGGCCCTGGCCCTGGCCGAGC 375
QY 381 GCTTCCACCGCTTCAGGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCGCCCA 440
Db 376 GCTTCCACCGCTTCAGGCCACCTATCCGTACCTGACGACGAGATCGACCTGCCGCCCA 435
QY 441 CCATCTCGTGTGACAGCGGGAGGAGCCCCACCTACAGGGCCCCCTGCACCCCTCCAGC 500
Db 436 CCATCTCGTGTGACAGCGGGAGGAGCCCCACCTACAGGGCCCCCTGCACCCCTCCAGC 495
QY 501 TTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGCGCACCCCAACA 560
Db 496 TTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGCGCACCCCAACA 555
QY 561 GAACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGCGGCCCTGCCGCCCA 620
Db 556 GAACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGCGGCCCTGCCGCCCA 615
QY 621 GCAGTAACCTCGGGCATCAGGCCACCTGCTACGGCAGCGGGCGGCATGGAGGGCGCGC 680
Db 616 GCAGTAACCTCGGGCATCAGGCCACCTGCTACGGCAGCGGGCGGCATGGAGGGCGCGC 675
QY 681 CGCCCACTACAGCGAGTTCATCGGCCACTACCGGGGTCTCTTCCAGCACGACGAGA 740
Db 676 CGCCCACTACAGCGAGTTCATCGGCCACTACCGGGGTCTCTTCCAGCACGACGAGA 735
QY 741 GCAGTGGCGGCCCTCTCTTGGTGGAGGGACCCGGTCCACCAACACACATCGCGCCCC 800
Db 736 GCAGTGGCGGCCCTCTCTTGGTGGAGGGACCCGGTCCACCAACACACATCGCGCCCC 795
QY 801 TAGAGAGCGCAGCCATCTCGAGCAAGAGAGGATAAACAGAAAAGGACACCCCTCTTAGG 860
Db 796 TAGAGAGCGCAGCCATCTCGAGCAAGAGAGGATAAACAGAAAAGGACACCCCTCTTAGG 855
QY 861 GTCCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGGCGAGAACTCCGCGCTTCT 920
Db 856 GTCCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGGCGAGAACTCCGCGCTTCT 915
QY 921 TAGAAGAGGAGTGAGAGGAGCGGGGGCGGCGCAGCAACCGCATCGTGTGGCCCTCCCTCC 980
Db 916 TAGAAGAGGAGTGAGAGGAGCGGGGGCGGCGCAGCAACCGCATCGTGTGGCCCTCCCTCC 975
QY 981 CACCTCCCTGTGTATAAATATTACATGTGTGTGTGAATGCACAAGCTAAGAGA 1040
Db 976 CACCTCCCTGTGTATAAATATTACATGTGTGTGTGAATGCACAAGCTAAGAGA 1035
QY 1041 GCTTGCAAAAAA 1057
Db 1036 GCTTGCAAAAAA 1052

RESULT 6
AAA75164
ID AAA75164 standard; cDNA; 969 BP.
XX
AC AAA75164;

AAA75165 standard; cDNA; 969 BP.

AAA75165;

15-JAN-2001 (first entry)

cDNA clone encoding a human TANGO 261 polypeptide.

TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
cellular proliferation; cellular differentiation; cellular adhesion;
von Willebrand factor-associated disorder; cell trafficking; cancer;
hematopoietic associated disease; atelectasis; pulmonary congestion;
oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
intestinal disorder; spleen associated disease; renal disorder;
cardiovascular disorder; ischemic heart disease; hydrocephalus;
brain herniation; iatrogenic disease; inflammation; meningitis;
Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 6..764

FT /*tag= a

FT /product= "TANGO 261"

XX WO200052022-A1.

PN

XX

PD 08-SEP-2000.

XX

PF 01-MAR-2000; 2000WO-US05226.

XX

PR 01-MAR-1999; 99US-0122458.

XX

PA (MILL-) MILLENNIUM PHARM INC.

XX

PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX

DR WPI; 2000-579269/54.

DR P-PSDB; AAB18463.

XX

PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -

XX

PS Disclosure; Page -; 175pp; English.

XX

CC AAA75163-65 encode human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.

CC note: the present sequence does not appear in the specification; it was
CC created using information provided.

XX

SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;

Query Match		85.8%;	Score 910;	DB 21;	Length 969;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 960;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	101	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTGTGATGATGATGATGATG	160		
DB	6	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTGTTGGTGTGATGATGATGATGATG	65		
QY	161	GTGGTGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATCAGCCGG	220		
DB	66	GTGGTGATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATCAGCCGG	125		
QY	221	CACAGCCAGGGCGGAGGAGAGAAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCG	280		
DB	126	CACAGCCAGGGCGGAGGAGAGAAAGATGCCCTGTCTCTCAGACGGATGCCTGTGGCCCTCG	185		
QY	281	GAGAGCACAGTGTCTAGGCAACGGAATCCAGAGCCCGCAGGTCTACGCCCCGCTCGGCCCC	340		
DB	186	GAGAGCACAGTGTCTAGGCAACGGAATCCAGAGCCCGCAGGTCTACGCCCCGCTCGGCCCC	245		
QY	341	ACCGACCGCTGGCCGTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCC	400		
DB	246	ACCGACCGCTGGCCGTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCC	305		
QY	401	ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTCTCAGACGGG	460		
DB	306	ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTCTCAGACGGG	365		
QY	461	GAGGAGCCCCACCTACCAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAG	520		
DB	366	GAGGAGCCCCACCTACCAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAG	425		
QY	521	CTGGAACCTGAACCGGGAGTCGGTCGCGGCACCCCCCAACAGAACCATCTTCGACAGTGAC	580		
DB	426	CTGGAACCTGAACCGGGAGTCGGTCGCGGCACCCCCCAACAGAACCATCTTCGACAGTGAC	485		
QY	581	CTGATGGATAGTGCCAGGCTGGGGCGGCCCTTGCCCCCCCAGCAGTAACTCGGGCATCAGC	640		
DB	486	CTGATGGATAGTGCCAGGCTGGGGCGGCCCTTGCCCCCCCAGCAGTAACTCGGGCATCAGC	545		
QY	641	GCCACGTGTACGGCAGCGGGCGGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTC	700		
DB	546	GCCACGTGTACGGCAGCGGGCGGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTC	605		
QY	701	ATCGGCCACTACCGGGGTCTCTCTTCCAGCACAGCAGAGCAGTGGGCGCGCTCCTTG	760		
DB	606	ATCGGCCACTACCGGGGTCTCTCTTCCAGCACAGCAGAGCAGTGGGCGCGCTCCTTG	665		
QY	761	CTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCATCTGG	820		
DB	666	CTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCATCTGG	725		
QY	821	AGCAAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCT	880		
DB	726	AGCAAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCT	785		
QY	881	GGGGTGTGAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAA	940		
DB	786	GGGGTGTGAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAA	845		
QY	941	GGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATA	1000		
DB	846	GGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATA	905		
QY	1001	TTTACATGTATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA	1060		
DB	906	TTTACATGTATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA	965		
QY	1061	A 1061			
DB	966	A 966			

[illegible]

Db	702	GCACAGTGTACAGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCCTCGGCCACCG	761
QY	345	ACCGCTGGCCGTGCCGCCCTTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCCCACCT	404
Db	762	ACCGCTGGCCGTGCCGCCCTTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCCCACCT	821
QY	405	ATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGCTGTCCAGCGGGAGG	464
Db	822	ATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGCTGTCCAGCGGGAGG	881
QY	465	AGCCCCACCCCTACAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG	524
Db	882	AGCCCCACCCCTACAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG	941
QY	525	AACCTGAACCGGAGTCGGTGCAGCACCCGCCACCCCAACAGAAACCATCTTCGACAGTGACCTGA	584
Db	942	AACCTGAACCGGAGTCGGTGCAGCACCCGCCACCCCAACAGAAACCATCTTCGACAGTGACCTGA	1001
QY	585	TGATAGTGCAGGCTGGCGGCCCTTGCCGCCCCAGCAGTAACCTCGGSCATCAGCGCCA	644
Db	1002	TGATAGTGCAGGCTGGCGGCCCTTGCCGCCCCAGCAGTAACCTCGGSCATCAGCGCCA	1061
QY	645	CGTGCTACGCGAGCGGGCGGCATGGAGGGCGGCCGCCACCTACAGCAGGTCTCATCG	704
Db	1062	CGTGCTACGCGAGCGGGCGGCATGGAGGGCGGCCGCCACCTACAGCAGGTCTCATCG	1121
QY	705	GCCACTACCGGGGTCTCTTCCAGCACACAGCAGCAGTGGGCCGCCCTCTTGTCTGG	764
Db	1122	GCCACTACCGGGGTCTCTTCCAGCACACAGCAGCAGTGGGCCGCCCTCTTGTCTGG	1181
QY	765	AGGGACCCGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA	824
Db	1182	AGGGACCCGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA	1241
QY	825	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGGGGGCGGGCTGGGG	884
Db	1242	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGGGGGCGGGCTGGGG	1301
QY	885	CTGCGTAGGTGAAAAGGCAG	904
Db	1302	CTGCGTAGGTGAAAAGGCAG	1321
RESULT 10			
AAI59654/c			
ID	AAI59654 standard; cDNA; 1069 BP.		
XX	AAI59654;		
XX	22-OCT-2001 (first entry)		
XX	Human polynucleotide SEQ ID NO 3643.		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
XX	leukaemia; ss.		
OS	Homo sapiens.		
XX	WO200153312-A1.		
PN	26-JUL-2001.		
XX	26-DEC-2000; 2000WO-US34263.		
PR	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		

CC or comprising an amino acid sequence encoded by one of 136 nucleotide
CC prostate specific nucleic acids sequences, PSNA, (or a sequence that
CC hybridises to it or is 60% identical to it), given in the specification.
CC Also included are a vector comprising the polynucleotide, a host cell
CC comprising the vector, an antibody specific for the PSP proteins and a
CC vaccine comprising the protein or polynucleotide. The PSP and PSNA are
CC useful for diagnosing and monitoring the presence and metastases of
CC prostate cancer in a patient. The PSNA is useful for determining the
CC the level PSNA in a sample. An antibody to the PSP is useful for
CC determining the presence of prostate specific protein in a sample, and
CC for treating a patient with prostate cancer, which induces an immune
CC response against the prostate cancer cell expressing the nucleic acid or
CC polypeptide and a kit is useful for detecting a risk of cancer or
CC presence of cancer in a patient. PSNA is useful as hybridisation probes
CC to detect, characterise and quantify hybridising nucleic acids from both
CC genomic and transcript-derived nucleic acid samples and also in
CC microarrays. Sequences of PSP and PSNA are useful as components in
CC databases for search analysis as well as in sequence analysis algorithms.
CC PSNA is useful to drive in vivo expression of PSP. The present
CC sequence is a PSNA of the invention.

XX
SQ Sequence 1583 BP; 361 A; 469 C; 457 G; 296 T; 0 other;

Query Match 48.0%; Score 509; DB 24; Length 1583;
Best Local Similarity 99.8%; Pred. No. 9.9e-194;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 313 GCCGCGGTCTACGCCCGCTCGGCCACCGACCGCTGGCGTCCGCTTCGCCCA 372
Db |||||
827 GCCGCGGTCTACGCCCGCTCGGCCACCGACCGCTGGCGTCCGCTTCGCCCA 886
Qy 373 GCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCAGATCGACCT 432
Db |||||
887 GCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCAGATCGACCT 946
Qy 433 GCGGCCACCATCTCGCTGTCAGACGGGAGGAGCCCCACCTACCAGGGCCCTGCAC 492
Db |||||
947 GCCACCCACCATCTCGCTGTCAGACGGGAGGAGCCCCACCTACCAGGGCCCTGCAC 1006
Qy 493 CCTCCAGTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGCGGCACC 552
Db |||||
1007 CCTCCAGTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGCGGCACC 1066
Qy 553 CCCAAACAGAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGCGGCCCTG 612
Db |||||
1067 CCCAAACAGAACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGCGGCCCTG 1126
Qy 613 CCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGCATGGA 672
Db |||||
1127 CCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGCATGGA 1186
Qy 673 GGGGCGCGCCACCTACAGCGAGTTCATCGGCCACTACCGGGGTCTCTCTCCAGCA 732
Db |||||
1187 GGGGCGCGCCACCTACAGCGAGTTCATCGGCCACTACCGGGGTCTCTCTCCAGCA 1246
Qy 733 CCAGCAGAGCAGTGGGCGCCCTCTGTGTGGAGGGGACCCGGCTCCACACACACAT 792
Db |||||
1247 CCAGCAGAGCAGTGGGCGCCCTCTGTGTGGAGGGGACCCGGCTCCACACACACAT 1306
Qy 793 CGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAAAGAGAAAGGACACCC 852
Db |||||
1307 CGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGAAAGAGAAAGGACACCC 1366
Qy 853 TCTCTAGGTCCTCCAGGGG 872
Db |||||
1367 TCTCTAGGTCCTCCAGGGG 1386

RESULT 12
AB284732/c
ID AB284732 standard; cDNA; 474 BP.
XX
AC AB284732;

XX
DT 14-MAY-2003 (first entry)
XX
DE Toxicologically relevant human nucleotide sequence #1891.
XX
KW Toxicologically relevant gene; toxicological response; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003016500-A2.
XX
PD 27-FEB-2003.
XX
PF 16-AUG-2002; 2002WO-US26514.
XX
PR 16-AUG-2001; 2001US-313080P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Neft RE, Dunn RT, Adkins K, Pickett GG, Kier LD, Schmeiser K;
PI Alen P;
XX
DR WPI; 2003-268322/26.
XX
PT Determining a toxicological response to an agent, useful for screening
PT of drugs, comprises comparing the expression profile of one or more
PT human toxic response genes to a reference gene expression profile
PT indicative of toxicity -
XX
PS Claim 1; Page 444; 455pp; English.

CC The present invention describes a method (M1) for determining a
CC toxicological response to an agent, which comprises comparing the
CC expression profile of one or more human toxic response genes to a
CC reference gene expression profile indicative of toxicity, and so
CC determining the presence of a toxic response to the agent. Also
CC described: (1) an array comprising one or more polynucleotides selected
CC from the genes corresponding to the partial sequences given in AB282842
CC to AB284764, or their fragments of at least 20 nucleotides, or
CC homologues; and (2) determining if a gene putatively identified to be a
CC toxic response gene plays a role on toxic response pathways by
CC determining the expression profile of the gene after exposure of cells
CC or a human subject to a known toxic pharmaceutical or industrial agent,
CC comprising: (a) exposing cells to an agent or isolating cells from a
CC human subject who was exposed to an agent; (b) obtaining the test gene
CC expression profile for a putatively identified toxic response gene after
CC exposure to a known toxic pharmaceutical or industrial agent; and
CC (c) comparing the test profile to the expression profile of a gene with
CC a similar function or comparing the test profile to the expression
CC profile of that gene after exposure to other known toxic compounds. The
CC methods are useful for predicting and determining toxicological responses
CC on a cellular, organ or system level. The arrays comprising the human
CC genes are useful for toxicological screening of drugs, pharmaceutical
CC compounds and chemicals.

XX
SQ Sequence 474 BP; 62 A; 136 C; 186 G; 90 T; 0 other;

Query Match 39.1%; Score 415; DB 25; Length 474;
Best Local Similarity 100.0%; Pred. No. 3.9e-156;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 297 GCAACGGAATCCCAGAGCGGAGGTCTACGCCCGCTCGGCCACCGCCTGGCGG 356
Db |||||
474 GCAACGGAATCCCAGAGCGGAGGTCTACGCCCGCTCGGCCACCGCCTGGCGG 415
Qy 357 TGC CGCCCTTCGCCCGCAGCGGAGCGCTTCCACCGCTTCCAGCCCATCTACCTATCCGTACCTGC 416
Db |||||
414 TGC CGCCCTTCGCCCGCAGCGGAGCGCTTCCACCGCTTCCAGCCCATCTACCTATCCGTACCTGC 355
Qy 417 AGCAGGAGATCGACCTGCCGCCACCGACCATCTCGCTGTGACAGCGGGAGGCCCCACCT 476
Db |||||
354 AGCAGGAGATCGACCTGCCGCCACCGACCATCTCGCTGTGACAGCGGGAGGCCCCACCT 295

QY 477 ACCAGGGCCCTCGACCCCTCCAGCTTCGGACCCCGAGCAGAGCTGGAACCGGG 536
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
294 ACCAGGGCCCTCGACCCCTCCAGCTTCGGACCCCGAGCAGAGCTGGAACCGGG 235
QY 537 AGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTCCA 596
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
234 AGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTCCA 175
QY 597 GGCTGGGGGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCA 656
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
174 GGCTGGGGGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCA 115
QY 657 GCGGCGGGCGCATGGAGGGGGCGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTA 711
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
114 GCGGCGGGCGCATGGAGGGGGCGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTA 60

RESULT 13
AAF65983
ID AAF65983 standard; cDNA; 408 BP.
XX
AC AAF65983;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1739.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-US18374.
PF
XX 02-JUL-1999; 99US-0142310.
PR
XX 02-JUL-1999; 99US-0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX

Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
XX Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 793; 1046pp; English.
XX

The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

SQ Sequence 408 BP; 81 A; 159 C; 115 G; 53 T; 0 other;
Query Match 33.0%; Score 350; DB 22; Length 408;
Best Local Similarity 99.8%; Pred. No. 3.5e-130;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 166 GATCACGTGCCTGCTGAGCCACTACAAGCTGTGTCACGGTCTTCATCAGCCGGCAGAG 225
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
8 GATCACGTGCCTGCTGAGCCACTACAAGCTGTGTCACGGTCTTCATCAGCCGGCAGAG 67
QY 226 CCAGGGCGGAGGAGAGAGATGCCCTGTCTCTAGAAGGATGCCCTGTGGCCCTCGAGAGAG 285
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
68 CCAGGGCGGAGGAGAGAGATGCCCTGTCTCTAGAAGGATGCCCTGTGGCCCTCGAGAGAG 127
QY 286 CACAGTGTACGGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCTTCGCCCCACCGA 345
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
128 CACAGTGTACGGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCTTCGCCCCACCGA 187
QY 346 CCGCCTGGCCGTGCCGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACTA 405
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
188 CCGCCTGGCCGTGCCGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCACTA 247
QY 406 TCCGTACCTGCAGACGAGATCGACCTGCCCGCCACACCATCTCGCTGTAGACGGGGAGGA 465
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
248 TCCGTACCTGCAGACGAGATCGACCTGCCACCCACCATCTCGCTGTAGACGGGGAGGA 307
QY 466 GCCCCCACCTACAGGGCCCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGAGCTGGA 525
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
308 GCCCCCACCTACAGGGCCCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGAGCTGGA 367
QY 526 ACTGAACCGGGAGTGGTGGCGCGCACCCCGCCCAACAGAACCA 566
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
368 ACTGAACCGGGAGTGGTGGCGCGCACCCCGCCCAACAGAACCA 408

RESULT 14
AAZ52964
ID AAZ52964 standard; cDNA; 812 BP.
XX
AC AAZ52964;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human prostate tumor cDNA library derived EST fragment #107.
XX
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
KW treatment; ds.
XX
OS Homo sapiens.
XX
PN DE19820190-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1998; 98DE-1020190.
XX
PR 28-APR-1998; 98DE-1020190.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-621386/54.
DR P-PSDB; AAY74135, AAY74136, AAY74137.
XX
PT New human nucleic acid sequences from pancreatic tumors, and related
XX proteins -
PS Claim 2; Page 269-270; 502pp; German.
XX
CC This invention describes novel polypeptides and their encoding nucleic
CC acids derived from human pancreatic tumor tissue which have cytostatic
CC activity. The sequences are also useful in producing pharmaceutical

CC compositions for treatment of pancreatic tumors. AAZ52858-253014
CC represent expressed sequence tag (EST) fragments derived from a human
CC pancreatic tumor cDNA library and which encode the proteins represented
CC in AAY73814-Y74252.
XX
SQ Sequence 812 BP; 157 A; 272 C; 237 G; 146 T; 0 other;
Query Match 30.3%; Score 322; DB 20; Length 812;
Best Local Similarity 100.0%; Pred. No. 5e-119;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 168 TCACGTGCTGCTGAGCCACTACAGCTGTCTGCACGCTCCTTCATCAGCCGCGCACAGCC 227
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 426 TCACGTGCTGCTGAGCCACTACAGCTGTCTGCACGCTCCTTCATCAGCCGCGCACAGCC 485
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 228 AGGGCGGAGGAGAGAAGATGCCCTGTCTCCTCAGAAGGATGCTGTGGCCCTCGGAGAGCA 287
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 486 AGGGCGGAGGAGAGAAGATGCCCTGTCTCCTCAGAAGGATGCTGTGGCCCTCGGAGAGCA 545
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 288 CAGTGTACGGCAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCCCTCGGCCACCGACC 347
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 546 CAGTGTACGGCAACGGAATCCAGAGCCGCGAGGTCTACGCCCGCCCTCGGCCACCGACC 605
QY 348 GCCTGGCGCTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACGCTTCCAGCCCATCTATC 407
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 606 GCCTGGCGCTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACGCTTCCAGCCCATCTATC 665
QY 408 CGTACTGCAGCAGCAGATCGACCTGCCGCCGCCACCATCTCGTGTGACAGCGGGAGGAGC 467
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 666 CGTACTGCAGCAGCAGATCGACCTGCCGCCGCCACCATCTCGTGTGACAGCGGGAGGAGC 725
QY 468 CCCACCTACCGAGGGCCCTG 489
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 726 CCCACCTACCGAGGGCCCTG 747

RESULT 15
ABK12143/c
ID ABK12143 standard; cDNA; 693 BP.
XX
AC ABK12143;
XX

DT 05-JUN-2002 (first entry)
XX
DE Human MIVR-1 homologous sequence #1.
XX
KW Human; ss; MIVR-1; Mechanically Induced Vascular Receptor 1;
KW cytosolic; cardiant; cerebroprotective; antiarteriosclerotic;
KW cardiac cell; anti-apoptotic; vascular endothelial cell;
KW cardiac hypertrophy; myocardial infarction; stroke; arteriosclerosis;
KW heart failure; AI761441.1.
XX
OS Homo sapiens.
XX
PN WO200216416-A2.
XX
PD 28-FEB-2002.
XX
PF 21-AUG-2001; 2001WO-US26089.
XX
PR 22-AUG-2000; 2000US-227159P.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
PA (PFIZ) PFIZER INC.
XX
PI Lee RT, Landschulz KT, Kennedy SP, Thompson JF, Turi TG;
XX
DR WPI; 2002-280912/32.
XX
PT Novel nucleic acid molecule encoding Mechanically Induced Vascular
PT Receptor-1 polypeptide, useful for treating cardiovascular diseases -
PS Disclosure. Page 101; 105pp; English.

XX The invention relates to an isolated nucleic acid molecule encoding a
CC Mechanically Induced Vascular Receptor (MIVR)-1 polypeptide having
CC cardiac cell anti-apoptotic activity and fragments of it provided
CC they are not identical to Genbank sequences AI761441.1, AI594390,
CC NM 004338 and AQ177461. Also included are expression vectors, host
CC cells, the MIVR-1 polypeptide, MIVR-1 binding peptides, modulators of
CC MIVR-1, contacting a molecule having cardiac cell anti-apoptotic activity
CC with a candidate agent, where the molecule is a nucleic acid molecule
CC comprising MIVR-1, IEX-1, VDUP-1, BTG-2 and TIS-11d or its
CC expression product, determining if the anti-apoptotic activity is
CC modulated and thereby identifying a modulator. The cardiac cell anti-
CC apoptotic molecules and nucleic acids of the invention are useful for
CC treating, diagnosing and monitoring progression of such diseases and
CC disorders as characterised by increased apoptotic cell-death of vascular
CC endothelial cells e.g. cardiac hypertrophy, myocardial infarction,
CC stroke, arteriosclerosis and heart failure. The present sequence
CC is one of the four Genbank sequences (AI761441.1) which are homologous to
CC the cDNA for human MIVR-1 and which are specifically disclaimed.
XX
SQ Sequence 693 BP; 101 A; 205 C; 237 G; 149 T; 1 other;
Query Match 29.5%; Score 313; DB 24; Length 693;
Best Local Similarity 99.3%; Pred. No. 2e-115;
Matches 583; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 468 CCCACCTACCGAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC 527
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 587 CCCACCTACCGAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC 528
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 528 TGAACCGGGAGTCGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGCATGATGG 587
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 527 TGAACCGGGAGTCGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGCATGATGG 468
QY 588 ATAGTGCCAGGCTGGCGGCCCTCCTGCCGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACGT 647
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 467 ATAGTGCCAGGCTGGCGGCCCTCCTGCCGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACGT 408
QY 648 GCTACGGCAGCGCGGCGCATGGAGGGGCGCGCCCGCCACCTACAGCGAGGTCTCGGCC 707
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 407 GCTACGGCAGCGCGGCGCATGGAGGGGCGCGCGCCCGCCACCTACAGCGAGGTCTCGGCC 348
QY 708 ACTACCGGGGTCCTCTTCCAGCACCATCATCGGCGCCCTCCTTCTGCTGGA-G 766
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 347 ACTACCGGGGTCCTCTTCCAGCACCATCATCGGCGCCCTCCTTCTGCTGAGG 288
QY 767 GGGACCGGCTCCACACACACATCGGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAA 826
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 287 GGGACCGGTTCCCGCACACACATCGGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAA 228
QY 827 GAGAAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGGCT 886
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 227 GAGAAAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGGGCTGGGGCT 168
QY 887 GCGTAGGTGAAAAGGAGAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGCGGG 946
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 167 GCGTAGGTGAAAAGGAGAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGCGGG 108
QY 947 GGGCGCAGCAACGCGATCGTGTGGCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCT 1006
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 107 GGGCGCAGCAACGCGATCGTGTGGCCCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCT 48
QY 1007 TGTGATGTCTGCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1053
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 47 TGTGATGTCTGCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1

Search completed: December 8, 2003, 04:24:34
Job time : 348 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 03:57:35 ; Search time 80 Seconds
(without alignments)
5853.844 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tctctctctgggttcgggtga.....cttgcaaaaaaaaaaaaaaa 1061

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

```

Word size :      0
Total number of hits satisfying chosen parameters: 1139956

```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #		DB	ID	Description
		Match	Length			
1	912	86.0	1140	4	US-09-769-482-1	Sequence 1, Appli
2	708	66.7	759	4	US-09-769-482-2	Sequence 2, Appli
3	24	2.3	24	4	US-09-769-482-5	Sequence 5, Appli
C 4	24	2.3	24	4	US-09-769-482-6	Sequence 6, Appli
C 5	24	2.3	24	4	US-09-769-482-10	Sequence 10, Appli
6	23	2.2	377	2	US-08-332-766A-1	Sequence 1, Appli
7	21	2.0	22	4	US-09-769-482-7	Sequence 7, Appli
8	21	2.0	686	2	US-08-522-421-7	Sequence 7, Appli
9	21	2.0	1327	1	US-08-165-315D-3	Sequence 3, Appli
C 10	21	2.0	8252	1	US-08-046-585-15	Sequence 15, Appli
C 11	21	2.0	8252	1	US-08-393-703-15	Sequence 15, Appli
C 12	21	2.0	8252	5	PCT-US93-11721-15	Sequence 15, Appli
C 13	21	2.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
14	20	1.9	20	4	US-09-769-482-9	Sequence 9, Appli
C 15	20	1.9	50	3	US-08-753-247-22	Sequence 22, Appli
16	20	1.9	51	3	US-08-753-247-23	Sequence 23, Appli
17	20	1.9	588	3	US-08-744-138-1	Sequence 1, Appli
18	20	1.9	588	4	US-09-241-376-1	Sequence 1, Appli
19	20	1.9	1189	1	US-07-781-034-4	Sequence 4, Appli
20	20	1.9	1189	5	PCT-US92-08328-4	Sequence 4, Appli
21	20	1.9	1218	4	US-09-685-853A-1	Sequence 1, Appli
22	20	1.9	1479	1	US-08-249-112-2	Sequence 2, Appli
23	20	1.9	1479	5	PCT-US95-06556-2	Sequence 2, Appli
24	20	1.9	1531	2	US-08-948-176-24	Sequence 24, Appli
25	20	1.9	2887	3	US-09-183-253-1	Sequence 1, Appli
26	20	1.9	319608	4	US-09-539-333D-1	Sequence 1, Appli
27	20	1.9	319608	4	US-09-679-409-1	Sequence 1, Appli

C	28	19	1.8	48	4	US-09-438-268-9	Sequence 9, Appl
	29	19	1.8	48	4	US-09-438-268-10	Sequence 10, Appl
C	30	19	1.8	49	1	US-08-155-171B-27	Sequence 27, Appl
	31	19	1.8	49	1	US-08-155-171B-28	Sequence 28, Appl
C	32	19	1.8	49	2	US-08-435-998-27	Sequence 27, Appl
	33	19	1.8	49	2	US-08-435-998-28	Sequence 28, Appl
	34	19	1.8	49	4	US-09-813-781-69	Sequence 69, Appl
C	35	19	1.8	51	4	US-09-438-268-11	Sequence 11, Appl
	36	19	1.8	51	4	US-09-438-268-12	Sequence 12, Appl
C	37	19	1.8	53	1	US-08-155-171B-36	Sequence 36, Appl
	38	19	1.8	53	2	US-08-435-998-36	Sequence 36, Appl
C	39	19	1.8	91	1	US-08-142-551B-129	Sequence 129, App
	40	19	1.8	91	1	US-08-142-551B-130	Sequence 130, App
	41	19	1.8	276	2	US-08-332-766A-30	Sequence 30, Appl
	42	19	1.8	586	3	US-09-385-982-139	Sequence 139, App
C	43	19	1.8	645	1	US-08-329-055-1	Sequence 1, Appl
	44	19	1.8	649	3	US-08-998-416-116	Sequence 116, App
	45	19	1.8	690	3	US-08-865-297-3	Sequence 3, Appl

ALIGNMENTS

```

RESULT 1
US-09-769-482-1
; Sequence 1, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1140
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(850)
US-09-769-482-1

```

Query Match	86.0%;	Score 912;	DB 4;	Length 1140;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 962; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
QY	95	CAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGCGTGATGATGGTG	154	
Dd	89	CAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGCGTGATGATGGTG	148	
QY	155	ATGGTGGTGGTGATCACGTGCCTGCTCAGCCACTACAAGCTGTCTGCACGGTCCCTTCATC	214	
Dd	149	ATGGTGGTGGTGATCACGTGCCTGCTCAGCCACTACAAGCTGTCTGCACGGTCCCTTCATC	208	
QY	215	AGCCGGCACAGCCAGGGGCGGAGGAGAAGATGCCCTGTCTCAGAAGGATGCCTGTGG	274	
Dd	209	AGCCGGCACAGCCAGGGGCGGAGGAGAAGATGCCCTGTCTCAGAAGGATGCCTGTGG	268	
QY	275	CCCTCGGAGAGCACAGTGTACGGCAACGGAATCCCAGAGCCGCAGGTCTACGCCCCGCCT	334	
Dd	269	CCCTCGGAGAGCACAGTGTACGGCAACGGAATCCCAGAGCCGCAGGTCTACGCCCCGCCT	328	
QY	335	CGGCCCAACCGACCGCCTGGCCGTGCCGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTC	394	

|||||
Db 329 CGGCCCCACGACCGCTGGCCCTGCGCCCTTCGCCACGCGGAGCGCTTCCACCGCTC 388
Qy 395 CAGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCGCCCCACCATCTCGCTGTCA 454
Db 389 CAGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCGCCCCACCATCTCGCTGTCA 448
Qy 455 GACGGGAGGAGCCCCACCTTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCGAG 514
Db 449 GACGGGAGGAGCCCCACCTTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCGAG 508
Qy 515 CAGCAGCTGGAACCTGAACCGGAGTCGGTGCGCGCACCCCAACAGAACCATCTTCGAC 574
Db 509 CAGCAGCTGGAACCTGAACCGGAGTCGGTGCGCGCACCCCAACAGAACCATCTTCGAC 568
Qy 575 AGTGACTGATGATAGTGCCAGGCTGGGCGGCCCTGCGCGGCCAGCAGTAACCTCGGC 634
Db 569 AGTGACTGATGATAGTGCCAGGCTGGGCGGCCCTGCGCGGCCAGCAGTAACCTCGGC 628
Qy 635 ATCAGCGCCACGTGCTACGGCAGCGGCGGCGCATGGAGGGCGCGCCGCCACCTACAGC 694
Db 629 ATCAGCGCCACGTGCTACGGCAGCGGCGGCGCATGGAGGGCGCGCCGCCACCTACAGC 688
Qy 695 GAGGTACCGGCCACTACCCGGGCTCTCTTCCAGCACACACATCGCGCCCTTAGAGAGCGAGCC 814
Db 749 TCCTTGTGGAGGGACCGGCTCCACACACACATCGCGCCCTTAGAGAGCGAGCC 808
Qy 815 ATCTGAGCAAGAGAGGATAAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGC 874
Db 809 ATCTGAGCAAGAGAGGATAAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGC 868
Qy 875 CGGGCTGGGGCTGCGTAGGTGAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGA 934
Db 869 CGGGCTGGGGCTGCGTAGGTGAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGA 928
Qy 935 GAGGAAGCGGGGGCGCAGCAGCATCGTGCGCCCTCCCTCCACCTCCCTGTGTA 994
Db 929 GAGGAAGCGGGGGCGCAGCAGCATCGTGCGCCCTCCCTCCACCTCCCTGTGTA 988
Qy 995 TAAATATTATCATGTGATGTCTGCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAA 1054
Db 989 TAAATATTATCATGTGATGTCTGCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAA 1048
Qy 1055 AAA 1057
Db 1049 AAA 1051

RESULT 2
US-09-769-482-2
; Sequence 2, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

; LENGTH: 759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-769-482-2
Query Match 66.7%; Score 708; DB 4; Length 759;
Best Local Similarity 99.9%; Pred. No. 1e-271;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 101 ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTATGATGATGATGATG 160
Db 1 ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTATGATGATGATGATG 60
Qy 161 GTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCAGGTCTTTCATCAGCGG 220
Db 61 GTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCAGGTCTTTCATCAGCGG 120
Qy 221 CACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGCGCCCTCG 280
Db 121 CACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGCGCCCTCG 180
Qy 281 GAGAGCACAGTGTACGCAACGGAATCCCAGAGCCGCGAGGTCTACGCCCGCCCTCGGCC 340
Db 181 GAGAGCACAGTGTACGCAACGGAATCCCAGAGCCGCGAGGTCTACGCCCGCCCTCGGCC 240
Qy 341 ACCGACCGCTGCGCGCTGCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCC 400
Db 241 ACCGACCGCTGCGCGCTGCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCAGCGCC 300
Qy 401 ACCTATCCGTACCTGCAGCACGAGATCGACCTGCGCCGCCACCATCTCGCTGTACAGCGG 460
Db 301 ACCTATCCGTACCTGCAGCACGAGATCGACCTGCGCCGCCACCATCTCGCTGTACAGCGG 360
Qy 461 GAGGAGCCCCCACCCTACAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 520
Db 361 GAGGAGCCCCCACCCTACAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 420
Qy 521 CTGGAACCTGAACCGGGAGTCGGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGA 580
Db 421 CTGGAACCTGAACCGGGAGTCGGTGCGCGCACCCCAACAGAACCATCTTCGACAGTGA 480
Qy 581 CTGATGGATAGTGCAGGCTGGGCGGCCCTTCCAGCACACAGAGAGAGAGAGAGAGAGAG 640
Db 481 CTGATGGATAGTGCAGGCTGGGCGGCCCTTCCAGCACACAGAGAGAGAGAGAGAGAGAG 540
Qy 641 GCCACGTGTACGGCAGCGCGCGCGCATGGAGGGGCGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTC 700
Db 541 GCCACGTGTACGGCAGCGCGCGCGCATGGAGGGGCGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTC 600
Qy 701 ATCGGCCACTACCGGGCTCCTCTTCCAGCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
Db 601 ATCGGCCACTACCGGGCTCCTCTTCCAGCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 761 CTGAGGGGACCGGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 820
Db 661 CTGAGGGGACCGGGCTCCACACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 720
Qy 821 AGCAAGAGAGAGATAAACAGAAAGGACACCCCTCTCTAG 859
Db 721 AGCAAGAGAGAGATAAACAGAAAGGACACCCCTCTCTAG 759

RESULT 3
US-09-769-482-5
; Sequence 5, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

;
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-5

Query Match 2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 900 GGCAGAACACTCCGCGCTTCTTAG 923
Db 1 GGCAGAACACTCCGCGCTTCTTAG 24

RESULT 4
US-09-769-482-6/c
; Sequence 6, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-6

Query Match 2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1022 GAATGCACAAGCTAAGAGAGCTTG 1045
Db 24 GAATGCACAAGCTAAGAGAGCTTG 1

RESULT 5
US-09-769-482-10/c
; Sequence 10, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED

;
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-10

Query Match 2.3%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1022 GAATGCACAAGCTAAGAGAGCTTG 1045
Db 24 GAATGCACAAGCTAAGAGAGCTTG 1

RESULT 6
US-08-332-766A-1
; Sequence 1, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-332-766A-1

Query Match 2.2%; Score 23; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 TGGTGGTGATGGTGATGGTG 160
|||||
Db 123 TGGTGGTGATGGTGATGGTG 145

RESULT 7
US-09-769-482-7
; Sequence 7, Application US/09769482
; Patent No. 6566130
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/09/769,482
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-7

Query Match 2.0%; Score 21; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CTTGGGTCGGGTGAAAGCGC 26
|||||
Db 1 CTTGGGTCGGGTGAAAGCGC 21

RESULT 8
US-08-522-421-7
; Sequence 7, Application US/08522421
; Patent No. 5908973
; GENERAL INFORMATION:
; APPLICANT: Abu-Bakar, Umi Kalsom
; APPLICANT: Barton, Sarah Louise
; APPLICANT: Gallego-Veigas, Pedro Pablo
; APPLICANT: Gray, Julie Elizabeth
; APPLICANT: Grierson, Donald
; APPLICANT: Lowe, Alexandra Louise
; APPLICANT: Picton, Steve
; APPLICANT: Whotton, Lee Colin
; TITLE OF INVENTION: DNA, DNA CONSTRUCTS, CELLS AND PLANTS
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/522,421
; FILING DATE: 11-JAN-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305868.3
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305869.1
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305859.2
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305866.7
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305867.5
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305860.0
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9305862.6
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9314351.9
; FILING DATE: 12-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9320988.0
; FILING DATE: 12-OCT-1993
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: ERT17
; US-08-522-421-7

Query Match 2.0%; Score 21; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1041 GCTTGCAAAAAAAAAAAAAA 1061
|||||
Db 663 GCTTGCAAAAAAAAAAAAAA 683

RESULT 9
US-08-165-315D-3
; Sequence 3, Application US/08165315D
; Patent No. 5525716
; GENERAL INFORMATION:
; APPLICANT: Odd-Arne Olsen
; APPLICANT: Roger Kalla
; TITLE OF INVENTION: Promoter
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: St. Onge, Steward, Johnston & Reens
; STREET: 986 Bedford Street
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06905
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC

OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/165,315D
FILING DATE: 10 December 1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9324707.0
FILING DATE: 2 December 1993
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: 2105-P0001
TELEPHONE: 201-324-6155
TELEFAX: 201-327-1096
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FRAGMENT TYPE: gene
FEATURE:
NAME/KEY: Ltp2 gene
US-08-165-315D-3

Query Match 2.0%; Score 21; DB 1; Length 1327;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 GTGATGATGGTGGTGGTG 163
|||
Db 918 GTGATGATGGTGGTGGTG 938

RESULT 10
US-08-046-585-15/c
Sequence 15, Application US/08046585
Patent No. 5453362
GENERAL INFORMATION:
APPLICANT: Lamarco, Kelly
APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
TITLE OF INVENTION: HOST CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/046,585
FILING DATE: 12-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57503-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 8252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-046-585-15

Query Match 2.0%; Score 21; DB 1; Length 8252;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 TGGTGGTGGTGGTGGTGG 158
|||
Db 2748 TGGTGGTGGTGGTGGTGG 2728

RESULT 11
US-08-393-703-15/c
Sequence 15, Application US/08393703
Patent No. 5585239
GENERAL INFORMATION:
APPLICANT: Lamarco, Kelly
APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
TITLE OF INVENTION: HOST CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,703
FILING DATE: 24-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57503-2/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 8252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-393-703-15

Query Match 2.0%; Score 21; DB 1; Length 8252;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 TGGTGGTGGTGGTGGTGG 158
|||
Db 2748 TGGTGGTGGTGGTGGTGG 2728

RESULT 12
PCT-US93-11721-15/c

Sequence 15, Application PC/TUS93111721
GENERAL INFORMATION:
APPLICANT: Lamarco, Kelly
APPLICANT: Wilson, Angus
APPLICANT: Herr, Winship
TITLE OF INVENTION: A NOVEL EKARYOTIC TRANSCRIPTION PROTEIN:
TITLE OF INVENTION: HOST CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11721
FILING DATE: 03-DEC-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: FP-57503-1/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 8252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
PCT-US93-11721-15

Query Match 2.0%; Score 21; DB 5; Length 8252;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 TGGTGGTGATGGTGATGG 158
Db 2748 TGGTGGTGATGGTGATGG 2728

RESULT 13
US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii
Patent No. 6503729
TITLE OF INVENTION: jannaschii
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii
FEATURE:
NAME/KEY: misc feature
LOCATION: (28222)..(28222)

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (103998)..(103998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600592)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature

LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
Query Match 2.0%; Score 21; DB 4; Length 1664976;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 138 TGGTGGTGATGATGGTGATGG 158
Db 568047 TGGTGGTGATGATGGTGATGG 568027
RESULT 14
US-09-769-482-9
Sequence 9, Application US/09769482
Patent No. 6566130
GENERAL INFORMATION:
APPLICANT: SRIVASTAVA, SHIV
APPLICANT: MOUL, JUDD W.
APPLICANT: XU, LINDA L.
APPLICANT: SEGAWA, TAKEHIKO
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
FILE REFERENCE: 04995.0057-00000
CURRENT APPLICATION NUMBER: US/09/769,482
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,772
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,045
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-769-482-9
Query Match 1.9%; Score 20; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 363 CCTTCGCCCCAGCGGAGCGC 382
Db 1 CCTTCGCCCCAGCGGAGCGC 20
RESULT 15
US-08-753-247-22/c
Sequence 22, Application US/08753247
Patent No. 6210929
GENERAL INFORMATION:
APPLICANT: SCHLOKAT, Uwe
APPLICANT: FISCHER, Bernhard
APPLICANT: FALKNER, Falko-Guenther
APPLICANT: DORNER, Friedrich
APPLICANT: BIBL, Johann
TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN
TITLE OF INVENTION: DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A
TITLE OF INVENTION: HETEROLOGOUS SEQUENCE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500

Search completed: December 8, 2003, 06:19:08
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 03:56:15 ; Search time 2638 seconds
(without alignments)
9775.227 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tctctcttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	722	68.0	729	13 BQ575741	BQ575741 UI-H-EZ1-
2	665	62.7	967	13 BQ641849	BQ641849 AGENCOURT
3	652	61.5	655	13 BQ691705	BQ691705 AGENCOURT
4	652	61.5	951	9 AL558881	AL558881

5	609	57.4	609	13	BQ636742	BQ636742	hd13h06.y
6	601	56.6	890	13	BQ690750	AGENCOURT	
C 7	599	56.5	602	14	CA431191	UI-H-FG1-	
C 8	585	55.1	588	13	BU624784	UI-H-FG1-	
C 9	570	53.7	629	13	BU730650	UI-E-C11-	
C 10	559	52.7	559	10	BE855409	7g13f05.x	
C 11	556	52.4	626	12	BM974296	UI-CF-EC1	
C 12	536	50.5	570	13	BQ575582	UI-H-EZ1-	
13	536	50.5	945	13	BU539219	AGENCOURT	
14	530	50.0	563	14	CB049800	NISC_gj13	
15	528	49.8	844	13	BQ686793	AGENCOURT	
16	528	49.8	952	13	BU157959	AGENCOURT	
C 17	523	49.3	730	12	BM677602	UI-E-EJ0-	
18	508	47.9	1046	12	BM922276	AGENCOURT	
19	505	47.6	646	13	BU859841	AGENCOURT	
20	499	47.0	551	12	BM141979	if25a11.y	
21	492	46.4	633	12	BM714472	UI-E-EJ0-	
C 22	491	46.3	782	12	BQ015170	UI-H-ED1-	
C 23	484	45.6	547	12	BM676516	UI-E-EJ0-	
24	484	45.6	552	12	BM713900	UI-E-EJ0-	
C 25	480	45.2	728	13	BU683523	UI-CF-EC1	
26	476	44.9	1007	9	AL558882		
27	469	44.2	613	10	BG680325	602629217	
28	468	44.1	964	13	BU859860	AGENCOURT	
29	461	43.4	461	12	BM712680	UI-E-EJ0-	
30	460	43.4	850	13	BU602918	AGENCOURT	
31	449	42.3	1280	13	BQ691500	AGENCOURT	
C 32	443	41.8	446	12	BM681946	UI-E-EJ0-	
33	442	41.7	973	13	BU169156	AGENCOURT	
34	441	41.6	668	14	CB044866	NISC_gc07	
35	441	41.6	916	13	BQ954555	AGENCOURT	
36	440	41.5	938	13	BU157842	AGENCOURT	
C 37	436	41.1	437	9	AI936228	wo63e04.x	
C 38	436	41.1	502	9	AI921394	wo24c07.x	
C 39	411	38.7	618	14	CD367193	UI-H-FT2-	
40	410	38.6	1127	13	BU174654	AGENCOURT	
C 41	402	37.9	1201	9	AL517150	AL517150	
C 42	399	37.6	451	9	AI493698	qy97c07.x	
C 43	392	36.9	874	13	BX362396	BX362396	
C 44	388	36.6	588	9	AI377498	tc37c07.x	
C 45	387	36.5	619	9	AI742327	wg50f07.x	

ALIGNMENTS

RESULT 1
BQ575741/c
LOCUS
DEFINITION
UI-H-EZ1-bbg-h-14-0-UI.s1 NCI CGAP_Ch2 Homo sapiens cDNA clone
UI-H-EZ1-bbg-h-14-0-UI 3', mRNA sequence.
ACCESSION
BQ575741
VERSION
BQ575741.1 GI:21479058
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 729)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopaedics
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD

FEATURES source Location/Qualifiers

1. 729

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-EZ1-bbg-h-14-0-UI"

/tissue_type="Chondrosarcoma Grade II"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP Ch2"

/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.

TAG_LIB=UI-H-EZ1

TAG_TISSUE=grade-2-chondrosarcoma

TAG_SEQ=ATCTAATATG

BASE COUNT 101 a 253 g 162 t

ORIGIN

Query Match 68.0%; Score 722; DB 13; Length 729;

Best Local Similarity 100.0%; Pred. No. 1.3e-150;

Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 GGCCACCGACCGCCTGGCCGTGCCGCTTCGCCAGCGGGAGCGCTTCCACCGCTTC 395

DB 729 GGCCACCGACCGCCTGGCCGTGCCGCTTCGCCAGCGGGAGCGCTTCCACCGCTTC 670

QY 396 AGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCACCATCTCGCTGTGAG 455

DB 669 AGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCACCATCTCGCTGTGAG 610

QY 456 ACGGGAGGAGCCCCACCTACCGAGCGCCCTGCACCTCCAGCTTCGGACCCCGAGC 515

DB 609 ACGGGAGGAGCCCCACCTACCGAGCGCCCTGCACCTCCAGCTTCGGACCCCGAGC 550

QY 516 AGCAGCTGGAAGTGAACCGGGAGTCGGTGGCGGCGACCCCCCAACAGAACCATCTTCGACA 575

DB 549 AGCAGCTGGAAGTGAACCGGGAGTCGGTGGCGGCGACCCCCCAACAGAACCATCTTCGACA 490

QY 576 GTGACCTGATGGATAGTGCCAGGCTGGCGGGCCCTGCCCCCAGCAGTAACCTCGGGCA 635

DB 489 GTGACCTGATGGATAGTGCCAGGCTGGCGGGCCCTGCCCCCAGCAGTAACCTCGGGCA 430

QY 636 TCAGCGCCAGCTGTACGGCAGCGGGCGGCATGGAGGGCGCGCCGCCACCTACAGCG 695

DB 429 TCAGCGCCAGCTGTACGGCAGCGGGCGGCATGGAGGGCGCGCCGCCACCTACAGCG 370

QY 696 AGGTATCGGCGCACTACCCGGGGTCTCTCTCCAGCACCGACAGCAGTGGCGCCCT 755

DB 369 AGGTATCGGCGCACTACCCGGGGTCTCTCTCCAGCACCGACAGCAGTGGCGCCCT 310

QY 756 CCTTGTGGAGGGACCCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCCA 815

DB 309 CCTTGTGGAGGGACCCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCCA 250

QY 816 TCTGGAGCAAGAGAGGATAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCC 875

DB 249 TCTGGAGCAAGAGAGGATAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCC 190

QY 876 GGGCTGGGCTGCGTAGGTGAAAGGAGAACCTCCGCGCTTCTTAGAAGAGGAGTGAG 935

Db 189 GGGCTGGGGCTGCGTAGTGAAGAGGAGCAACACTCGCGCTTCTTAGAAGAGAGTGAG 130

QY 936 AGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTAT 995

Db 129 AGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTAT 70

QY 996 AAATATTACATGTGTGTGTGTGAATGCACACAGCTAAGAGAGCTTGCAAAAAAAA 1055

Db 69 AAATATTACATGTGTGTGTGTGAATGCACACAGCTAAGAGAGCTTGCAAAAAAAA 10

QY 1056 AA 1057

Db 9 AA 8

RESULT 2

BQ641849

LOCUS

DEFINITION BQ641849 967 bp mRNA linear EST 15-JUL-2002

5', mRNA sequence.

ACCESSION BQ641849

VERSION BQ641849.1 GI:21766021

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 967)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LCM2493 row: 9 column: 18

High quality sequence stop: 571.

Location/Qualifiers

1. 967

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6292265"

/tissue_type="normal pigmented retinal epithelium"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 43"

/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library. |"

BASE COUNT 194 a 334 c 296 g 143 t

ORIGIN

Query Match 62.7%; Score 665; DB 13; Length 967;

Best Local Similarity 99.9%; Pred. No. 3.8e-138;

Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 107 GAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTATGATGGTGGTGGTGGTGGT 166

Db 1 GAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 60

QY 167 ATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGACAGC 226

Db 61 ATCAGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGACAGC 120

Qy 227 CAGGGCGGAGGAGAGAGATGCCCTGTCTCCTCAGAAAGGATGCTGTGGCCCTCGGAGAGC 286
Db 121 CAGGGCGGAGGAGAGAGATGCCCTGTCTCCTCAGAAAGGATGCTGTGGCCCTCGGAGAGC 180
Qy 287 ACAGTGTAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCCACCCGAC 346
Db 181 ACAGTGTAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCCACCCGAC 240
Qy 347 CGCTGGCCGTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTAT 406
Db 241 CGCCTGGCCGTGCCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTAT 300
Qy 407 CCCTACTGCAGCACGAGATCGACCTGCCGCCCCACCATCTCGTGTTCAGACGGGGAGG 466
Db 301 CCGTACTGCAGCACGAGATCGACCTGCCGCCCCACCATCTCGTGTTCAGACGGGGAGG 360
Qy 467 CCCCCACCTACCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGTGGAA 526
Db 361 CCCCCACCTACCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGTGGAA 420
Qy 527 CTGAACCGGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGCCTGATG 586
Db 421 CTGAACCGGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGCCTGATG 480
Qy 587 GATAGTGCAGGCTGGCGGCCCTGCCGCCCCAGCAGTAACCTTCGACAGTGCCTGATG 646
Db 481 GATAGTGCAGGCTGGCGGCCCTGCCGCCCCAGCAGTAACCTTCGAGCGCCAGC 540
Qy 647 TGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCCCACTACAGCGAGGTCTCGGC 706
Db 541 TGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCCCACTACAGCGAGGTCTCGGC 600
Qy 707 CACTACCCGGGTCTCTTCAGACACAGCAGCAGTGGCGGCCCTCCTTGTGCTGGAG 766
Db 601 CACTACCCGGGTCTCTTCAGACACAGCAGCAGTGGCGGCCCTCCTTGTGCTGGAG 660
Qy 767 GGGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAG 822
Db 661 GGGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAG 716

RESULT 3
BQ691705
LOCUS
DEFINITION BQ691705 655 bp mRNA linear EST 15-JUL-2002
5', mRNA sequence.
ACCESSION BQ691705.1 GI:21817021
VERSION BQ691705.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 655).
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LCM2367 row: g column: 20
High quality sequence stop: 645.
Location/Qualifiers
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source

/db_xref="taxon:9606"
/clone="IMAGE:6208939"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 155 a 216 c 193 g 91 t
ORIGIN
Query Match 61.5%; Score 652; DB 13; Length 655;
Best Local Similarity 100.0%; Pred. No. 4e-135;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 410 TACCTGCAGCAGAGATCGACCTGCCGCCCCACCATCTCGCTGTTCAGACGGGGAGGAGCCC 469
Db 1 TACCTGCAGCAGAGATCGACCTGCCGCCCCACCATCTCGCTGTTCAGACGGGGAGGAGCCC 60
Qy 470 CCACCCCTACAGGGCCCTTCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAATG 529
Db 61 CCACCCCTACAGGGCCCTTCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAATG 120
Qy 530 AACCGGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGCCTGATGGAT 589
Db 121 AACCGGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGCCTGATGGAT 180
Qy 590 AGTGCCAGGCTGGCGGGCCCTTGCCTCCAGCAGTAACCTCGGGCATCAGCGCCACGTGC 649
Db 181 AGTGCCAGGCTGGCGGGCCCTTGCCTCCAGCAGTAACCTCGGGCATCAGCGCCACGTGC 240
Qy 650 TACGGCAGCGCGGGCGCATGGAGGGCGCGCCCACTACAGCAGGTTCATCGGCCAC 709
Db 241 TACGGCAGCGCGGGCGCATGGAGGGCGCGCCCACTACAGCAGGTTCATCGGCCAC 300
Qy 710 TACCCGGGTCTCTCTCCAGCACAGCAGCAGTGGCGGCCCTCTCTTGTGAGGGG 769
Db 301 TACCCGGGTCTCTCTCCAGCACAGCAGCAGTGGCGGCCCTCTCTTGTGAGGGG 360
Qy 770 ACCCGGTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAG 829
Db 361 ACCCGGTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAG 420
Qy 830 AAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGGGCTGGGCTGCG 889
Db 421 AAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGGGCTGGGCTGCG 480
Qy 890 TAGGTGAAAAGGCAGAACACTCCGGCTTCTTTAGAGAGGAGTGAGAGGAAGCGGGGGG 949
Db 481 TAGGTGAAAAGGCAGAACACTCCGGCTTCTTTAGAGAGGAGTGAGAGGAAGCGGGGGG 540
Qy 950 CGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTTACATGT 1009
Db 541 CGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTTACATGT 600
Qy 1010 GATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAAAGAGAGAGAGAG 1061
Db 601 GATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAAAGAGAGAGAGAG 652

RESULT 4

AL558881/c
LOCUS
DEFINITION AL558881 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ015YF12 3-PRIME, mRNA sequence.
ACCESSION AL558881
VERSION AL558881.2 GI:31283014
KEYWORDS EST.

AL558881 951 bp mRNA linear EST 31-MAY-2003
AL558881 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ015YF12 3-PRIME, mRNA sequence.
AL558881
AL558881.2 GI:31283014
EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 951)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12903836.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9945.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015DC06NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015DC06NP1.
FEATURES
Location/Qualifiers
1..951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015VF12"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 141 a 275 c 318 g 201 t 16 others
ORIGIN
Query Match 61.5%; Score 652; DB 9; Length 951;
Best Local Similarity 99.6%; Pred. No. 2.9e-135;
Matches 802; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 168 TCACGTGCTGCTGAGCCACTACAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCC 227
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
846 TCACGTGCTGCTGAGCCACTACAGCTGTCTGCACGGTCTTCATCAGCCGGAACAGCC 787
Qy 228 AGGGGGGAGGAGAGAGATGCCCTGTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCA 287
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
786 AGGGGGGAGGAGAGAGATGCCCTGTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCA 727
Qy 288 CAGTGTCAAGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCCTCGGCCACCGACC 347
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
726 CAGTGTCAAGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCCTCGGCCACCGACC 667
Qy 348 GCCTGGCCGTGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATC 407
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
666 GCCTGGCCGTGCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATC 607
Qy 408 CGTACCTGCAGCAGAGATCGACCTTCCGCCCCACCATCTCGCTGTTCAGACGGGAGGAGC 467
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
606 CGTACCTGCAGCAGAGATCGACCTTCCGCCCCACCTTCTCGCTTTCAGACGGGAGGAGC 547
Qy 468 CCCACCTACAGGGCCCTTGACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC 527
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
546 CCCACCTACAGGGCCCTTGACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC 487
Qy 528 TGAACCGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGG 587
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
486 TGAACCGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGG 427
Qy 588 ATAGTGCCAGGCTGGGGCGCCCTGCCCGCCCGCAGCAGTAACCTCGGGCATCAGCGCCACGT 647
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
426 ATAGTGCCAGGCTGGGGCGCCCTGCCCGCCCGCAGCAGTAACCTCGGGCATCAGCGCCACGT 367

Qy 648 GCTACGGCAGCGCGGCGCATGGAGGGCGCGCCACCTACAGCGAGGTTCATCGGCC 707
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
366 GCTACGGCAGCGCGGCGCATGGAGGGCGCGCCACCTACAGCGAGGTTCATCGGCC 307
Qy 708 ACTACCCGGGTCTCTTCCAGCACACAGCAGCAGAGCAGTGGGGCGCCCTCTTGTGGAGG 767
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
306 ACTACCCGGGTCTCTTCCAGCACACAGCAGCAGCAGCAGTGGGGCGCCCTCTTGTGGAGG 247
Qy 768 GGACCCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAG 827
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
246 GGACCCGGCTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAG 187
Qy 828 AGAAGGATAAACAGAAAGACACACCTCTCTAGGGTCCCAGGGGGCGGGCTGGGCTG 887
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
186 AGAAGGATAAACAGAAAGACACACCTCTCTAGGGTCCCAGGGGGCGGGCTGGGCTG 127
Qy 888 CGTAGGTGAAAAGGCGAGCACTCCCGCTTCTTAGAAGGAGTGAGAGGAAGCGCGGG 947
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
126 CGTAGGTGAAAAGGCGAGCACTCCCGCTTCTTAGAAGGAGTGAGAGGAAGCGCGGG 67
Qy 948 GCGCGAGCAACGCATCGTGTGGCCC 972
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
66 GCGCGAGCAACGCATCGTGTGGCCC 42

RESULT 5
BQ636742
LOCUS
DEFINITION BQ636742 609 bp mRNA linear EST 15-JUL-2002
hd13h06.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he
Homo sapiens cDNA clone hd13h06 5', mRNA sequence.
ACCESSION BQ636742
VERSION BQ636742.1 GI:21761201
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 609)
Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A., Touchman
J.W., Bouffard, G., Smith, D. and Peterson, K.
TITLE Expressed sequence tag analysis of human retina for the NEIBank
Project: Retbindin, an abundant, novel retinal cDNA and alternative
splicing of other retina-preferred gene transcripts
JOURNAL Mol. Vis. 8 (4), (2002) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 13 row: h column: 06
Seq primer: M13RP1 reverse primer (ABI).

FEATURES
Location/Qualifiers
1..609
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hd13h06"
/tissue_type="Retina"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retina cDNA (Un-normalized, unamplified
): hd/he"
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue
was dissected from two 80 year old donors with no observed
eye disease. 100ug of total RNA was used for library
construction. A directionally cloned cDNA library in the
pSPORT1 vector (Life Technologies) was constructed at
Bioserve Biotechnology (Laurel MD) essentially following
the protocols of the SuperScript Plasmid System full
details of which are contained in the manufacturer's

Instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor (5'-pGACTAGTTCTAGATCGGAGCGCGCC(T)15-3'). EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

BASE COUNT 114 a 238 c 182 g 75 t

Query Match 57.4%; Score 609; DB 13; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.3e-125;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	214	CAGCCGGCACAGCCAGGGCGGAGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTG	273
Db	1	CAGCCGGCACAGCCAGGGCGGAGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTG	60
Qy	274	GCCCTCGGAGAGACAGTGTCAAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCC	333
Db	61	GCCCTCGGAGAGACAGTGTCAAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCC	120
Qy	334	TCGGCCACCCAGCCGCTGGCCGTGCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTT	393
Db	121	TCGGCCACCCAGCCGCTGGCCGTGCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTT	180
Qy	394	CCAGCCACCTATCCGTACCTGACGACGAGATCGACCTGCGCCACCATCTCGCTGTC	453
Db	181	CCAGCCACCTATCCGTACCTGACGACGAGATCGACCTGCGCCACCATCTCGCTGTC	240
Qy	454	AGACGGGAGGAGCCCCACCTACAGGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGA	513
Db	241	AGACGGGAGGAGCCCCACCTACAGGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGA	300
Qy	514	GCAGCAGCTGGAAGTGAACCGGGAGTGGGTGCGCGCACCCCAACAGAACCATCTTCGA	573
Db	301	GCAGCAGCTGGAAGTGAACCGGGAGTGGGTGCGCGCACCCCAACAGAACCATCTTCGA	360
Qy	574	CAGTGACCTGATGGATAGTCCAGGCTGGCGGCCCTGCCCCCAGCAGTAACCTCGG	633
Db	361	CAGTGACCTGATGGATAGTCCAGGCTGGCGGCCCTGCCCCCAGCAGTAACCTCGG	420
Qy	634	CATCAGCCACGCTGCTACGGCAGCGGGCGGCATGGAGGGGGCGCCACCTACAG	693
Db	421	CATCAGCCACGCTGCTACGGCAGCGGGCGGCATGGAGGGGGCGCCACCTACAG	480
Qy	694	CGAGTCTCGGCCACTACCCGGGTCTCTCTCCAGCACACAGAGCAGTGGGCCGCC	753
Db	481	CGAGTCTCGGCCACTACCCGGGTCTCTCTCCAGCACACAGAGCAGTGGGCCGCC	540
Qy	754	CTCCTTGCTGGAGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGC	813
Db	541	CTCCTTGCTGGAGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGC	600
Qy	814	CATCTGGAG 822	
Db	601	CATCTGGAG 609	

RESULT 6
BQ690750
LOCUS BQ690750 890 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8046394 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6209341
5', mRNA sequence.

ACCESSION BQ690750
VERSION BQ690750.1 GI:21816066
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2368 row: h column: 14
High quality sequence stop: 627.
Location/Qualifiers

FEATURES

source

1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209341"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 168 a 303 c 273 g 146 t

ORIGIN

Query Match 56.6%; Score 601; DB 13; Length 890;
Best Local Similarity 99.8%; Pred. No. 5.5e-124;
Matches 651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	TCCTCCTTGGGTTCCGGTGAAAGCGCTTGGGGTTTCAGTGGCCCATGATCCCCGAGCTGC	60
Db	24	TCCTCCTTGGGTTCCGGTGAAAGCGCTTGGGGTTTCAGTGGCCCATGATCCCCGAGCTGC	83
Qy	61	TGGAGAACTGAAGGGGACGGTCTCTCGGAAACCAGGCAATGGCGAGCTGGAGTTGT	120
Db	84	TGGAGAACTGAAGGGGACGGTCTCTCGGAAACCAGGCAATGGCGAGCTGGAGTTGT	143
Qy	121	TCAGATCATCATCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	180
Db	144	TCAGATCATCATCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	203
Qy	181	GAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAG	240
Db	204	GAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAG	263
Qy	241	AGAAGATGCCCTGTCTCAGAAGGATGCCTGTGGCCCTCGAGAGACACAGTGTCAAGCAA	300
Db	264	AGAAGATGCCCTGTCTCAGAAGGATGCCTGTGGCCCTCGAGAGACACAGTGTCAAGCAA	323
Qy	301	CGGAATCCAGAGCCGAGTCTACGCCCGCCCTCGGCCACCGCCCTGGCCGTGCCGTGCC	360
Db	324	CGGAATCCAGAGCCGAGTCTACGCCCGCCCTCGGCCACCGCCCTGGCCGTGCCGTGCC	383
Qy	361	GCCCTTGGCCCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCA	420
Db	384	GCCCTTGGCCCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCA	443
Qy	421	CGAGATCGACCTGCCGCCACCATCTCGCTGTCTCAGACGGGAGGAGCCCCACCTACCA	480
Db	444	CGAGATCGACCTGCCGCCACCATCTCGCTGTCTCAGACGGGAGGAGCCCCACCTACCA	503
Qy	481	GGGCCCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGCAGTGGAACTGAACCGGGAGTC	540
Db	504	GGGCCCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGCAGTGGAACTGAACCGGGAGTC	563
Qy	541	GGTGGCGCACCCCCAAACAGAACCATCTTCGACAGTGCATGATGATGATGATGATGATGATG	600

```
Db 564 GGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATGATGCCAGGCT 623
Qy 601 GGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTAC 652
Db 624 GGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTAC 675

RESULT 7
CA431191/c
LOCUS
DEFINITION
  UI-H-FG1-bgi-c-12-0-UI.s1 linear EST 07-NOV-2002
  UI-H-FG1-bgi-c-12-0-UI 3', mRNA sequence.
ACCESSION
  CA431191
VERSION
  CA431191.1 GI:24793917
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 602)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: James Martin
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@uiowa.edu
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  source
    Location/Qualifiers
      1..602
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="UI-H-FG1-bgi-c-12-0-UI"
        /tissue_type="Cell lines"
        /dev_stage="Adult"
        /lab_host="DH10B (Life Technologies)"
        /clone_lib="NCI CGAP FG1"
        /note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
        with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
        NCI CGAP_FG1 is a normalized cDNA library obtained from a
        pool of mRNA from 2 cell lines from Enchondroma tissues.
        The library was constructed according to Bonaldo, Lennon
        and Soares, Genome Research, 6:791-806, 1996. First strand
        cDNA synthesis was primed with an oligo-dT primer
        containing a Not I site. Double stranded cDNA was ligated
        to an EcoR I adaptor, digested with Not I, and cloned
        directionally into pT7T3-Pac vector. The oligonucleotide
        used to prime the synthesis of first-strand cDNA contains
        a library tag sequence that is located between the Not I
        site and the (dT)18 tail. The sequence tag for this
        library is CGGTCACTC. The cell lines were provided by Dr.
        James Martin from the University of Iowa.
        TAG LIB=UI-H-FG1
        TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
        TAG_SEQ=CGGTCACTC"
BASE COUNT      83 a 179 c 196 g 144 t
ORIGIN

Query Match      56.5%; Score 599; DB 14; Length 602;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 463 GGAGCCCCACCCCTACCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCT 522
Db 602 GGAGCCCCACCCCTACCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCT 543
```

```
Qy 523 GGAACCTGAACCGGGAGTCGGTGC GCGCACCCCAACAGAACCATCTTCGACAGTGACCT 582
Db 542 GGAACCTGAACCGGGAGTCGGTGC GCGCACCCCAACAGAACCATCTTCGACAGTGACCT 483
Qy 583 GATGATAGTCCAGGCTGGGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGC 642
Db 482 GATGATAGTCCAGGCTGGGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGC 423
Qy 643 CACGTGCTACGGCAGCGCGGGCGCATGGAGGGGCGGCCGCCACCTACAGCGAGGTGAT 702
Db 422 CACGTGCTACGGCAGCGCGGGCGCATGGAGGGGCGGCCGCCACCTACAGCGAGGTGAT 363
Qy 703 CGGCCACTACCCGGGGTCTCTCTCCAGCACCAAGCAGTGGGCCCGCCCTCCTTGCT 762
Db 362 CGGCCACTACCCGGGGTCTCTCTCCAGCACCAAGCAGTGGGCCCGCCCTCCTTGCT 303
Qy 763 GGAGGGGACCCGGTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAG 822
Db 302 GGAGGGGACCCGGTCCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAG 243
Qy 823 CAAAGAGAGAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGGGCTGG 882
Db 242 CAAAGAGAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGGGCTGG 183
Qy 883 GGCTGCGTAGGTGAAAGCGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAG 942
Db 182 GGCTGCGTAGGTGAAAGCGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAG 123
Qy 943 CGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATATT 1002
Db 122 CGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAATATT 63
Qy 1003 TACATGTGATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1061
Db 62 TACATGTGATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 4

RESULT 8
BU624784/c
LOCUS
DEFINITION
  UI-H-FG1-bgi-f-22-0-UI.s1 NCI CGAP_FG1 Homo sapiens cDNA clone
  UI-H-FG1-bgi-f-22-0-UI 3', mRNA sequence.
ACCESSION
  BU624784
VERSION
  BU624784.1 GI:23290999
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 588)
REFERENCE
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: James Martin
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@uiowa.edu
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  source
    Location/Qualifiers
      1..588
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="UI-H-FG1-bgi-f-22-0-UI"
        /tissue_type="Cell lines"
        /dev_stage="Adult"
        /lab_host="DH10B (Life Technologies)"
```

```

/clone_lib="NCI_CGAP_FG1"
/note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_FG1 is a normalized cDNA library obtained from a
pool of mRNA from 2 cell lines from Enchondroma tissues.
The library was constructed according to Bonaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is CGGTCACTC. The cell lines were provided by Dr.
James Martin from the University of Iowa.
TAG_LIB=UI-H-FG1
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG_SEQ=CGGTCACTC"

```

```
Query Match          55.1%; Score 585; DB 13; Length 588;
Best Local Similarity 100.0%; Pred. No. 2.7e-120;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	477	ACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCCGGG	536
Db	588	ACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCCGGG	529
QY	537	AGTCGGTGC GCGCACCCCCAAACAGAAACCATCTTCGACAGTGACCTGATGGATAGTGCCA	596
Db	528	AGTCGGTGC GCGCACCCCCAAACAGAAACCATCTTCGACAGTGACCTGATGGATAGTGCCA	469
QY	597	GGCTGGGCGGCCCTGCCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCA	656
Db	468	GGCTGGGCGGCCCTGCCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCA	409
QY	657	GCGGCGGGCGCATGGAGGGCGCGCGCCACCTACAGCGAGGTCACTCGGCCACTACCCGG	716
Db	408	GCGGCGGGCGCATGGAGGGCGCGCGCCACCTACAGCGAGGTCACTCGGCCACTACCCGG	349
QY	717	GGTCCTCCTTCCAGCACAGCAGAGCAGTGGCGCGCCCTCCTTCTGTGGAGGGGACCCCGC	776
Db	348	GGTCCTCCTTCCAGCACAGCAGAGCAGTGGCGCGCCCTCCTTCTGTGGAGGGGACCCCGC	289
QY	777	TCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCCATCTGGAGCAAGAGAGGATA	836
Db	288	TCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCCATCTGGAGCAAGAGAGGATA	229
QY	837	AACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGGCCGGCTGGGGCTGCGTAGGTGA	896
Db	228	AACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGGCCGGCTGGGGCTGCGTAGGTGA	169
QY	897	AAAGGCAGAACACTCCGCGTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCA	956
Db	168	AAAGGCAGAACACTCCGCGTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCA	109
QY	957	ACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGCT	1011
Db	108	ACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGCT	49
QY	1017	GGTCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAAAAAAAAAA	1061
Db	48	GGTCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAAAAAAAAAA	4

RESULT 9	
BU730650/c	
LOCUS	629 bp mRNA
DEFINITION	UI-E-C11-afq-o-09-0-UI.s1 UI-E-C11 Homo sapiens cDNA clone
ACCESSION	UI-E-C11-afq-o-09-0-UI 3', mRNA sequence.
	BU730650

VERSION	REFERENCE	JOURNAL
KEYWORDS	AUTHORS	MEDLINE
SOURCE	TITLE	PUBMED
ORGANISM		COMMENT

BUT30650.1 GI:23654753
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 629)
 Bonaldo,M.F.; Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
source

```

Location/Qualifiers
1. .629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afq-o-09-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI).
TAG_LIB=UI-E-C11
TAG_TISSUE=RPE and Choroid
TAG_SEQ=ACCTA"

```

BASE COUNT	87 a	187 c	206 g	148 t	1 others
ORIGIN					

Query Match 53.7%; Score 570; DB 13; Length 629;
Best Local Similarity 100.0%; Pred. No. 5.2e-117;
Matches 570; Conservative 0; Mismatches 0; Indels

Qy	488	TGCACCCCTCCAGCTTCGGGACCCCAGCAGCAGACTTGGAACTGAACCGGGAGTCGGTGCGC	547
Db	577	TGCACCCCTCCAGCTTCGGGACCCCAGCAGCAGACTTGGAACTGAACCGGGAGTCGGTGCGC	518
Qy	548	GCACCCCCCAAAACAGAACCATCTTCGACAGTAGCCTGTATGGATAGTGCCAGGCTGGGGCGGC	607
Db	517	GCACCCCCCAAAACAGAACCATCTTCGACAGTAGCCTGTATGGATAGTGCCAGGCTGGGGCGGC	458
Qy	608	CCCTGCCCCCCCAGCAGTAACCTCGGGCATTCAGCGCCAAGTCTACGGCACGGCGGGGCGC	667


```
|||||
Db 457 CCTGCCCCCCCAGCAGTAACCTCGGCATCAGCGCCACGTGCTACGGCAGCGCGGCGC 398
Qy 668 ATGAGGGGGCGCGCCACCTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTTC 727
Db 397 ATGAGGGGGCGCGCCACCTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTTC 338
Qy 728 CAGCACCAGCAGCAGTGCGGCCCTCTTGTGAGGGGACCGGCTCCACCACACA 787
Db 337 CAGCACCAGCAGCAGTGCGGCCCTCTTGTGAGGGGACCGGCTCCACCACACA 278
Qy 788 CACATCGCGCCCTTAGAGAGCGCAGCCCATCTGGAGCAAAGAGAGATAAACAGAAAGGA 847
Db 277 CACATCGCGCCCTTAGAGAGCGCAGCCCATCTGGAGCAAAGAGAGATAAACAGAAAGGA 218
Qy 848 CACCTCTCTAGGTTCCCGAGGGGGCGGGCTGGGGCTGGTAGGTGAAAAGGCAGAAC 907
Db 217 CACCTCTCTAGGTTCCCGAGGGGGCGGGCTGGGGCTGGTAGGTGAAAAGGCAGAAC 158
Qy 908 ACTCGCGCTTCTTAGAAGAGAGAGTAGAGGAAGCGGGGGCGGCAGCAACGCATCGT 967
Db 157 ACTCGCGCTTCTTAGAAGAGAGTAGAGGAAGCGGGGGCGGCAGCAACGCATCGT 98
Qy 968 GGCCCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGCTGGTCTGAATGC 1027
Db 97 GGCCCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGCTGGTCTGAATGC 38
Qy 1028 ACAAGCTAAGAGAGCTTGCAAAAAAAAAA 1057
Db 37 ACAAGCTAAGAGAGCTTGCAAAAAAAAAA 8
```

```
RESULT 10
BE855409/c
LOCUS
DEFINITION
7913f05.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3306369 3'
similar to TR:Q9UJD3 Q9UJD3 DJ718J7.1 ;, mRNA sequence.
ACCESSION
BE855409
VERSION
BE855409.1 GI:10367404
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 559)
AUTHORS
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 443.
FEATURES
Location/Qualifiers
1..559
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3306369"
/tissue_type="glioblastoma (pooled)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn23"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
```

```
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 82 a 173 c 182 g 122 t
ORIGIN
Query Match 52.7%; Score 559; DB 10; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.5e-114;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 488 TGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGC 547
Db 559 TGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACTGAACCGGGAGTCGGTGC 500
Qy 548 GCACCCCCAAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGCGGC 607
Db 499 GCACCCCCAAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGCGGC 440
Qy 608 CCCTGCCCCCCCAGCAAGTAACCTCGGGCATCAGCGGCACAGTCTACGGCAGCGGGCGGC 667
Db 439 CCCTGCCCCCCCAGCAAGTAACCTCGGGCATCAGCGGCACAGTCTACGGCAGCGGGCGGC 380
Qy 668 ATGAGGGGGCGCGCCACCTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTCTTC 727
Db 379 ATGAGGGGGCGCGCCACCTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTCTTC 320
Qy 728 CAGCACCAGCAGCAGTGCGGGCGCCCTCTCTTGTGAGGGGACCCCGCTCCACCACACA 787
Db 319 CAGCACCAGCAGCAGTGCGGGCGCCCTCTCTTGTGAGGGGACCCCGCTCCACCACACA 260
Qy 788 CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAGAGATAAACAGAAAGGA 847
Db 259 CACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAGAGATAAACAGAAAGGA 200
Qy 848 CACCTCTCTAGGGTCCCAGGGGGCGCGGGTGGGGTGGCTAGGTGAAAGGCGAGAAC 907
Db 199 CACCTCTCTAGGGTCCCAGGGGGCGCGGGTGGGGTGGCTAGGTGAAAGGCGAGAAC 140
Qy 908 ACTCGCGCTTCTTAGAAGAGAGTAGAGAGGAAGCGGGGGCGGCAGCAACGCATCGT 967
Db 139 ACTCGCGCTTCTTAGAAGAGAGTAGAGAGGAAGCGGGGGCGGCAGCAACGCATCGT 80
Qy 968 GGCCCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGCTGGTCTGAATGC 1027
Db 79 GGCCCTCCCTCCACCTCCCTGTGTATAAATATTTACATGTGATGCTGGTCTGAATGC 20
Qy 1028 ACAAGCTAAGAGAGCTTGC 1046
Db 19 ACAAGCTAAGAGAGCTTGC 1
```

```
RESULT 11
BM974296/c
LOCUS
DEFINITION
BM974296 626 bp mRNA linear EST 20-FEB-2003
UI-CF-EC1-aca-k-23-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
UI-CF-EC1-aca-k-23-0-UI 3', mRNA sequence.
ACCESSION
BM974296
VERSION
BM974296.1 GI:19591887
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 626)
AUTHORS
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
```


MEDLINE	97044477	
PUBMED	8889548	
COMMENT	Contact: McCray, PB	
	McCray Lab	
	University of Iowa	
	2024 University of Iowa Med Labs, Iowa City, IA 52242, USA	
	Tel: 319 356 4866	
	Fax: 319 356 7171	
	Email: paul-mccray@uiowa.edu	
	Tissue Procurement: Dr. M. J. Welsh, University of Iowa	
	CDNA Library preparation: Dr. M. Bento Soares, University of Iowa	
	CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa	
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
	Clone Distribution: Researchers may obtain clones from Research	
	Genetics (www.resgen.com) or from Open Biosystems	
	(www.openbiosystems.com).	
	Seq primer: M13 FORWARD	

FEATURES	source
Location/Qualifiers	
1. .626	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="UI-CF-EC1-aca-k-23-0-UI"	
/tissue_type="Lung"	
/dev_stage="Adult and Fetal"	
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"	
/clone_lib="UI-CF-EC1"	
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AAGTGCTTAC."	
TAG_LIB=UI-CF-EC1	
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-377 and 380-383	
TAG_SEQ=AAGTGCTTAC"	
85 a	187 c 201 g 151 t 2 others
BASE COUNT	
ORIGIN	

RESULT 12
BQ575582/c
LOCUS
DEFINITION
BQ575582
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BQ575582 570 bp mRNA linear EST 19-JUN-2002
UI-H-EZ1-bbf-f-09-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone
UI-H-EZ1-bbf-f-09-0-UI 3', mRNA sequence.
BQ575582
BQ575582.1 GI:21478899
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopaedics
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES	Location/Qualifiers	source
	1..570	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="UI-H-EZ1-bbf-f-09-0-UI"	
	/tissue_type="Chondrosarcoma Grade II"	
	/dev_stage="Adult"	
	/lab_host="DH10B (Life Technologies)"	
	/clone_lib="NCI CGAP Ch2"	
	/note="Organ: Left pelvis; Vector: pTT3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is	

TGATCACGCT.
TAG LIB=UI-H-E21
TAG TISSUE=grade-2-chondrosarcoma
TAG_SEQ=ATCTAATATG"
BASE COUNT 80 a 174 c 177 g 139 t
ORIGIN
Query Match 50.5%; Score 536; DB 13; Length 570;
Best Local Similarity 100.0%; Pred. No. 1.8e-109;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 522 TGGAACTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACC 581
Db 543 TGGAACTGAACCGGAGTGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACC 484
QY 582 TGATGGATAGTCCAGGCTGGGGCGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGG 641
Db 483 TGATGGATAGTCCAGGCTGGGGCGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGG 424
QY 642 CCACGTGTACGGCAGCGCGGGCGGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCA 701
Db 423 CCACGTGTACGGCAGCGCGGGCGGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCA 364
QY 702 TCGGCCACTACCCGGGTCTCTTCAGCACCACATCGCGCCCTAGAGAGCGCAGCCATCTGGA 761
Db 363 TCGGCCACTACCCGGGTCTCTTCAGCACCACATCGCGCCCTAGAGAGCGCAGCCATCTGGA 304
QY 762 TGGAGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGA 821
Db 303 TGGAGGGACCCGGCTCCACCACACATCGCGCCCTAGAGAGCGCAGCCATCTGGA 244
QY 822 GCAAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGGCTG 881
Db 243 GCAAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGGCTG 184
QY 882 GGGCTGCTAGGTGAAAGGAGCACTCCGGCTCTCTAGAGAGGAGTGAGAGGAAG 941
Db 183 GGGCTGCTAGGTGAAAGGAGCACTCCGCGCTCTCTAGAGAGGAGTGAGAGGAAG 124
QY 942 GCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATAT 1001
Db 123 GCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATAT 64
QY 1002 TTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1057
Db 63 TTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 8

RESULT 13
BU539219 945 bp mRNA linear EST 13-SEP-2002
LOCUS
DEFINITION
AGENCOURT_10215265 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6569922 5', mRNA sequence.
ACCESSION
BU539219
VERSION
BU539219.1 GI:22849660
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 945)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LCM2757 row: p column: 18
High quality sequence stop: 663.
Location/Qualifiers
1..945
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6569922"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 194 a 319 c 287 g 144 t 1 others
ORIGIN
Query Match 50.5%; Score 536; DB 13; Length 945;
Best Local Similarity 99.5%; Pred. No. 1.2e-109;
Matches 736; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 18 TGAAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGCTGGAGAACTGAAGCGG 77
Db 1 TGAAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGCTGGAGAACTGAAGCGG 60
QY 78 ACGGTCTCTTCGAAACCAAGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCG 137
Db 61 ACAGTCTCTTCGAAACCAAGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCG 120
QY 138 TGGTGGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 197
Db 121 TGGTGGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
QY 198 CTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCT 257
Db 181 CTGCACGGTCTCTTCATCAGTGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCT 240
QY 258 CAGAAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCG 317
Db 241 CAGAAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCG 300
QY 318 AGGTCTACGCCCGCTCGGCCACCGACCGCTGGCGTGGCGCCCTTCGCCCGAGCGG 377
Db 301 AGGTCTACGCCCGCTCGGCCACCGACCGCTGGCGTGGCGCCCTTCGCCCGAGCGG 360
QY 378 AGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGACGACGAGATCGACCTGCCG 437
Db 361 AGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGACGACGAGATCGACCTGCCG 420
QY 438 CCACCATCTCGCTGTACAGCGGGGAGGAGCCCCACCTACAGGGCCCCCTGCACCTCC 497
Db 421 CCACCATCTCGCTGTACAGCGGGGAGGAGCCCCACCTACAGGGCCCCCTGCACCTCC 480
QY 498 AGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGACCCCAAA 557
Db 481 AGCTTCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGACCCCAAA 540
QY 558 ACAGAAACCATCTTCGACAGTACCTGATGGATAGTGCAGAGTGGCGGCCCTGCCCCC 617
Db 541 ACAGAAACCATCTTCGACAGTACCTGATGGATAGTGCAGAGTGGCGGCCCTGCCCCC 600
QY 618 CCAGCAGTAACCTCGGGCATCAGCGCCACGTGTACGGCAGCGGGCGGCGCATGGAGGGG 677
Db 601 CCAGCAGTAACCTCGGGCATCAGCGCCACGTGTACGGCAGCGGGCGGCGCATGGAGGGG 660
QY 678 CGCGCCCCACCTACAGCGAGGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACCG 737
Db 661 CGCGCCCCACCTACAGCGAGGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACCG 720

QY 738 AGAGCAGTGGCGCCCTCC 757
|||||
Db 721 AGAGCAGTGGCGCCCTCC 740
|||||
RESULT 14
CB049800 563 bp mRNA linear EST 17-JAN-2003,
LOCUS NISC.g113c01.y1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3271656
DEFINITION 5', mRNA sequence.
ACCESSION CB049800
VERSION CB049800.1 GI:27788087
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 563)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
cDNA Library Preparation:
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLAM8009 row: E column: 1
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3271656"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 146 a 170 c 168 g 79 t
ORIGIN
Query Match 50.0%; Score 530; DB 14; Length 563;
Best Local Similarity 100.0%; Pred. No. 3.9e-108;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 528 TGAACCGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGACCTGATGG 587
|||||
Db 10 TGAACCGGAGTCGGTGGCGGACCCCAACAGAACCATCTTCGACAGTGACCTGATGG 69
|||||
QY 588 ATAGTGCCAGGCTGGGCGGCCCTGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACGT 647
|||||
Db 70 ATAGTGCCAGGCTGGGCGGCCCTGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACGT 129
|||||
QY 648 GCTACGGCAGCGCGGCGGCATGAGGGGCGCGCCGCCACCTACAGCGAGTCAATCGGCC 707
|||||
Db 130 GCTACGGCAGCGCGGCGGCATGAGGGGCGCGCCGCCACCTACAGCGAGTCAATCGGCC 189
|||||
QY 708 ACTACCGGGGTCTCTCTCCAGCACCAAGAGAGAGTGGGCGCCCTCTCTGCTGAGG 767
|||||

Db 190 ACTACCGGGGTCTCTCTCCAGCACACAGAGAGTGGGCGCCCTCTCTGCTGAGG 249
|||||
QY 768 GGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAG 827
|||||
Db 250 GGACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAG 309
|||||
QY 828 AGAAGGATAAACAGAAAGACACCCCTCTCTAGGGTCCCGGGGGCGGCTGGGGCTG 887
|||||
Db 310 AGAAGGATAAACAGAAAGACACCCCTCTCTAGGGTCCCGGGGGCGGCTGGGGCTG 369
|||||
QY 888 CGTAGGTGAAAGGCGAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGG 947
|||||
Db 370 CGTAGGTGAAAGGCGAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGG 429
|||||
QY 948 GCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACAT 1007
|||||
Db 430 GCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACAT 489
|||||
QY 1008 GTGATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1057
|||||
Db 490 GTGATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 539
|||||
RESULT 15
BQ686793 844 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8345390 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248044
DEFINITION 5', mRNA sequence.
ACCESSION BQ686793
VERSION BQ686793.1 GI:21812109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 844)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2387 row: e column: 05
High quality sequence stop: 593.
Location/Qualifiers
1..844
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6248044"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 239 a 236 c 235 g 133 t 1 others
ORIGIN
Query Match 49.8%; Score 528; DB 13; Length 844;
Best Local Similarity 99.8%; Pred. No. 7.6e-108;

Matches 578;		Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	483	GCCCTGACACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGG 542							
Db	1								
		1 GCCCTGACACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGG 60							
Qy	543	TGGCGGACACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGG 602							
Db	61								
		61 TGGCGGACACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGG 120							
Qy	603	GGGCCCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCACCTGCTACGGCAGCGGCG 662							
Db	121								
		121 GCGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCACCTGCTACGGCAGCGGCG 180							
Qy	663	GGCGATGGAGGGGCGCGCCCACTACAGCGAGGTTCATCGGCCACTACCCGGGGTCCT 722							
Db	181								
		181 GCGCATGGAGGGGCGCGCCCACTACAGCGAGGTTCATCGGCCACTACCCGGGGTCCT 240							
Qy	723	CCTTCCAGCACAGCAGAGCAGTGGGCGCCCTCCTTGCTGGAGGGGACCCGGCTCCACC 782							
Db	241								
		241 CCTTCCAGCACAGCAGAGCAGTGGGCGCCCTCCTTGCTGGAGGGGACCCGGCTCCACC 300							
Qy	783	ACACACACATCGGCGCCCTAGAGAGCGCGAGCCATCTGGAGCAAGAGAGGATAAACAGA 842							
Db	301								
		301 ACACACACATCGGCGCCCTAGAGAGCGCGAGCCATCTGGAGCAAGAGAGGATAAACAGA 360							
Qy	843	AAGGACACCTCTCTAGGGTCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGGC 902							
Db	361								
		361 AAGGACACCTCTCTAGGGTCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGGC 420							
Qy	903	AGAACTCCGCGCTTCTTAGAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCAT 962							
Db	421								
		421 AGAACTCCGCGCTTCTTANAAGAGAGTGAGAGGAAGCGGGGGCGCAGCAACGCAT 480							
Qy	963	CGTGTGGCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGCTGTCTG 1022							
Db	481								
		481 CGTGTGGCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGCTGTCTG 540							
Qy	1023	AATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1061							
Db	541								
		541 AATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 579							

Search completed: December 8, 2003, 06:17:44
Job time : 2648 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 21:17:14 ; Search time 344 Seconds
(without alignments)
8325.883 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tcctccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq 19Jun03.*	
				1:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.*
				2:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.*
				3:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.*
				4:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.*
				5:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.*
				6:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.*
				7:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.*
				8:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.*
				9:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.*
				10:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.*
				11:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.*
				12:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.*
				13:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.*
				14:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.*
				15:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.*
				16:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.*
				17:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.*
				18:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.*
				19:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.*
				20:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.*
				21:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.*
				22:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.*
				23:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.*
				24:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.*
				25:	/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1061	100.0	1061	21	Sequence encoding
2	1057.8	99.7	1334	25	Human secretory po
3	1034.4	97.5	1140	24	Prostate cancer-as
4	1021.2	96.2	1066	22	Human polynucleoti
5	991.2	93.4	1069	22	Human polynucleoti
6	961.8	90.7	969	21	CDNA encoding a hu
7	960.2	90.5	969	21	CDNA clone encodin
8	960.2	90.5	969	21	CDNA clone encodin

9	960.2	90.5	969	21	AAA75165	CDNA clone encodin
10	800.6	75.5	1321	24	ABK12137	Human CDNA encodin
11	728.8	68.7	1583	24	ABS61424	Prostate specific
12	644.2	60.7	693	24	ABK12143	Human MIVR-1 homol
13	523.8	49.4	878	24	ABK12142	Mouse cDNA encodin
14	521.4	49.1	1713	21	AAA75152	CDNA encoding a mu
15	519.8	49.0	1713	21	AAA75166	CDNA clone encodin
16	519.8	49.0	1713	21	AAA75167	CDNA clone encodin
17	519.8	49.0	1713	21	AAA75168	CDNA clone encodin
18	477.8	45.0	1879	23	AAS84503	DNA encoding novel
19	418.6	39.5	474	25	ABZ84732	Toxicologically re
20	399.8	37.7	408	22	AAF65983	Novel human polynu
21	399	37.6	673	24	ABT09178	Phase-1 Rat Ct gen
22	365.6	34.5	812	20	AAS252964	Human prostate tum
23	352.4	33.2	937	21	AAS252507	Human secreted pro
24	350.8	33.1	920	24	ABK34251	Human CDNA for nov
25	350	33.0	8093	24	ABK12145	Human MIVR-1 homol
26	349	32.9	8494	23	AAS77304	DNA encoding novel
27	345.8	32.6	8065	19	AAV38335	Manic-depressive i
28	334.8	31.6	8103	23	AAS77312	DNA encoding novel
29	281	26.5	8440	23	AAS77305	DNA encoding novel
30	281	26.5	8440	24	ABK83477	Human CDNA differe
31	277.8	26.2	8011	19	AAV38336	Manic-depressive i
32	253.6	23.9	8041	23	AAS77310	DNA encoding novel
33	249.2	23.5	475	24	ABK12144	Human MIVR-1 homol
34	239.4	22.6	426	23	AAS84502	DNA encoding novel
35	232	21.9	254	21	AAA41265	Human secreted exp
36	223.4	21.1	1428	23	AAS92497	DNA encoding novel
37	223.4	21.1	2942	23	AAS77313	DNA encoding novel
38	188.4	17.8	522	24	ABT10027	Human breast cance
39	60	5.7	60	24	ABN40872	Human spliced tran
40	52.8	5.0	114955	20	AAX53491	Human adenosine A1
41	51	4.8	51	22	AAH89714	Human coding sequ
42	48.4	4.6	1337	20	AAS17263	Human gene express
43	48.4	4.6	11304	22	AAL03113	Human reproductive
44	48.2	4.5	336	24	ABS54598	DNA encoding varia
45	48	4.5	114955	20	AAX53491	Human adenosine A1

ALIGNMENTS

RESULT 1

AAA47429
ID AAA47429 standard; DNA; 1061 BP.
XX
AC AAA47429;
XX
DT 20-OCT-2000 (first entry)
XX
DE Sequence encoding human neuron-associated protein.

XX
KW Neuron associated protein; NEUP; neurological disorder; epilepsy;
KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;
KW Alzheimer's disease; Pick's disease; Huntington's disease;
KW dementia; Parkinson's disease; demyelinating disease; meningitis;
KW prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;
KW cerebral palsy; muscular dystrophy; central nervous system; CNS;
KW peripheral nervous system; PNS; myopathy; schizophrenia;
KW actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;
KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;
KW AIDS; Addison's disease; adult respiratory distress syndrome;
KW allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
KW Werner syndrome, trauma; human; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 101..859

/tag= a

/product= Neuron associated protein

XX AC ABK92120;
XX DT 15-AUG-2002 (first entry)
XX DE Prostate cancer-associated DNA sequence #6.
XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX KW gene therapy; gene; ds.
XX OS Mammalia.
XX PN WO200230268-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US32045.
XX PR 13-OCT-2000; 2000US-0687576.
XX PR 08-DEC-2000; 2000US-0733288.
XX PR 08-DEC-2000; 2000US-0733742.
XX PR 24-JAN-2001; 2001US-263957P.
XX PR 16-MAR-2001; 2001US-276791P.
XX PR 16-MAR-2001; 2001US-276888P.
XX PR 06-APR-2001; 2001US-281922P.
XX PR 24-APR-2001; 2001US-286214P.
XX PR 30-APR-2001; 2001US-0847046.
XX PR 04-MAY-2001; 2001US-288589P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
XX DR P-PSDB; ABG61805.
XX PT Detecting a prostate cancer-associated transcript in a cell in a
XX PT patient, useful for diagnosing prostate cancer (PC) or screening
XX PT modulators of PC, by determining if prostate cancer-associated genes
XX PT are expressed in a prostate tissue -
XX PS Claim 22; Page 305; 436pp; English.
XX CC The present invention relates to methods of detecting a prostate
XX CC cancer-associated transcript in a cell from a patient. The method
XX CC comprises contacting a biological sample from the patient with
XX CC prostate cancer-associated polynucleotides (designated PC genes) that
XX CC selectively hybridise to a sequence that is at least 80% identical
XX CC to them. The prostate cancer-associated polynucleotide sequences
XX CC are differentially expressed in prostate tumour tissue or in
XX CC prostate cancer and are derived from the tissues of various
XX CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
XX CC The methods of the invention are useful for diagnosing and treating
XX CC prostate cancer in mammals. The prostate cancer-associated genes are
XX CC useful for diagnosing or treating prostate cancer, as well as for
XX CC identifying modulators of prostate cancer or agents that inhibit
XX CC prostate cancer. The nucleic acid sequences are particularly useful
XX CC in gene therapy, as a vaccine or in antisense applications.
XX CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
XX CC sequences.
XX SQ Sequence 1140 BP; 270 A; 350 C; 336 G; 184 T; 0 other;
Query Match 97.5%; Score 1034.4; DB 24; Length 1140;
Best Local Similarity 99.2%; Pred. No. 2.6e-214;
Matches 1050; Conservative 0; Mismatches 6; Indels 2; Gaps 1;
Qy 4 TCCTTGGGTTGGGTGAAGCGCTTGGGGTTTCAGTGGGGCCATGATCCCCGAGCTGCTGG 63
Dd 1 TCCTTGGGTTGGGTGAAGCGCTTGGGGTTTCAGTGGGGCCATGATCCCCGAGCTGCTGG 58
Qy 64 AGAAGTGAAGCGCGGACGGTCTCTCTGCGAAACCCAGGCAATGGCGAGCTGGAGTTTGTCA 123

Db 59 AGAACTGAAGGCGGACAGTCTCTCGAAACCAGGCAATGGCGGAGCTGGAGTTTGTCA 118
Qy 124 GATCATCATCGTGGTGGTGATGATGGTGGTGATGGTGGTGATCACGTGCCTGCTGAG 183
Dd 119 GATCATCATCGTGGTGGTGATGATGGTGGTGATGGTGGTGATCACGTGCCTGCTGAG 178
Qy 184 CCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCCGGCACAGCCGGGGGAGGAGAGA 243
Dd 179 CCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCCGGCACAGCCGGGGGAGGAGAGA 238
Qy 244 AGATGCCCTGTCTCAGAAGGATGCTTGGCCCTCGGAGAGCACAGTGTACGGCAACGG 303
Dd 239 AGATGCCCTGTCTCAGAAGGATGCTTGGCCCTCGGAGAGCACAGTGTACGGCAACGG 298
Qy 304 AATCCAGAGCCGCGAGGTCTACGCCCGCTCGGCCACCGACCGCTGGCCGTGCCGCC 363
Dd 299 AATCCAGAGCCGCGAGGTCTACGCCCGCTCGGCCACCGACCGCTGGCCGTGCCGCC 358
Qy 364 CTTGCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGACGA 423
Dd 359 CTTGCGCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGACGA 418
Qy 424 GATCGACCTGCCGCCACCATCTCGTGTGACAGCGGGGAGGAGCCCCACCTACCGGG 483
Dd 419 GATCGACCTGCCGCCACCATCTCGTGTGACAGCGGGGAGGAGCCCCACCTACCGGG 478
Qy 484 CCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTGGT 543
Dd 479 CCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTGGT 538
Qy 544 GCGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGATGATGCCAGGTGGG 603
Dd 539 GCGCGCACCCCAACAGAACCATCTTCGACAGTACCTGATGATGATGCCAGGTGGG 598
Qy 604 CGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGG 663
Dd 599 CGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGCGG 658
Qy 664 GCGCATGGAGGGGCGCGCCACCTACAGCGAGGTATCGGCCACTACCCGGGGTCTCTC 723
Dd 659 GCGCATGGAGGGGCGCGCCACCTACAGCGAGGTATCGGCCACTACCCGGGGTCTCTC 718
Qy 724 CTTCCAGCACCGCAGCAGTGGGCGCCCTCTTCTGGAGGGGACCCGGCTCCACCA 783
Dd 719 CTTCCAGCACCGCAGCAGTGGGCGCCCTCTTCTGGAGGGGACCCGGCTCCACCA 778
Qy 784 CACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAA 843
Dd 779 CACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAA 838
Qy 844 AGGACACCTCTCTAGGTTCCCGAGGGGGCGCGGCTGGGGCTCGTAGGTGAAAAGGCA 903
Dd 839 AGGACACCTCTCTAGGTTCCCGAGGGGGCGCGGCTGGGGCTCGTAGGTGAAAAGGCA 898
Qy 904 GAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 963
Dd 899 GAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
Qy 964 GTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTACATGTGATGTCTGTCTGA 1023
Dd 959 GTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTACATGTGATGTCTGTCTGA 1018
Qy 1024 ATGCACAAGCTAAGAGAGCTTGCAAAAAA 1061
Dd 1019 ATGCACAAGCTAAGAGAGCTTGCAAAAAA 1056
RESULT 4
AAI57868
ID AAI57868 standard; cDNA; 1066 BP.
XX
AC AAI57868;
XX

DT	22-OCT-2001	(first entry)
XX		
DE	Human polynucleotide SEQ ID NO 71.	
XX		
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	leukaemia; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200153312-A1.	
XX		
PD	26-JUL-2001.	
XX		
PF	26-DEC-2000; 2000WO-US34263.	
XX		
PR	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
PR	09-JUL-2000; 2000US-0598042.	
PR	19-JUL-2000; 2000US-0620312.	
PR	03-AUG-2000; 2000US-0653450.	
PR	14-SEP-2000; 2000US-0662191.	
PR	19-OCT-2000; 2000US-0693036.	
PR	29-NOV-2000; 2000US-0727344.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;	
XX		
DR	WPI; 2001-442253/47.	
DR	P-PSDB; AAM38712.	
XX		
PT	Novel nucleic acids and polypeptides, useful for treating disorders	
PT	such as central nervous system injuries -	
XX		
PS	Claim 1; SEQ ID NO 71; 10078pp; English.	
XX		
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and	
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,	
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful	
CC	in gene therapy. A composition containing a polypeptide or polynucleotide	
CC	of the invention may be used to treat diseases of the peripheral nervous	
CC	system, such as peripheral nervous injuries, peripheral neuropathy and	
CC	localised neuropathies and central nervous system diseases, such as	
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
CC	utilisation of the activities such as: Immune system suppression,	
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
CC	assays for receptor activity, arthritis and inflammation, leukaemias and	
CC	C.N.S disorders.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification.	
XX		
SQ	Sequence 1066 BP; 225 A; 348 C; 325 G; 168 T; 0 other;	
	Query Match	96.2%; Score 1021.2; DB 22; Length 1066;
	Best Local Similarity	98.6%; Pred. No. 1.8e-211;
	Matches 1043; Conservative	0; Mismatches 8; Indels 7; Gaps 1;
Qy	1	TCCTCCTGGGTTCCGGTGAAGCGCTTGGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
Dd	16	TCCTCCTGGGTTCCGGTGAAGCGCTTGGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 75
Qy	61	TGGAGAACTGAAGCGGACGGTCTCCTGCGAAACCAAGGCAATGGCGGAGCTGGAGTTTGT 120
Dd	76	TGGAGAACTGAAGCGGACAGTCTCCTGCGAAACCA-----GCGGACGACGATCTTGT 128

DE Human polynucleotide SEQ ID NO 3643.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

PN 26-JUL-2001.

PD 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

DR P-PSDB; AAM40498.

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Claim 1; SEQ ID NO 3643; 10078pp; English.

PS The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX the encoded polypeptides (AAM38642-AAM42213) with neurotropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX Sequence 1069 BP; 176 A; 323 C; 346 G; 222 T; 2 other;

SQ Query Match 93.4%; Score 991.2; DB 22; Length 1069;

Best Local Similarity 97.8%; Pred. No. 5.7e-205;

Matches 1037; Conservative 0; Mismatches 15; Indels 8; Gaps 3;

QY 1 TCCTCCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60

DB 1054 TCCTCCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 995

QY 61 TGGAGAACTGAAGCGGACGGTCTCTCGGAAACCAGGCAATGGCGGAGCTGGAGTTGT 120

DB 994 TGGAGAACTGAAGCGGACAGTCTCTCGGAAACCAGCGGAGCTGAAG-----ATNTTGT 941

QY 121 TCAGATCATCATCAT-CGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 179

Db 940 TCAGATCATCATCATCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 881

QY 180 TGAGCCACTACAAGCTGTCTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGGCGGAGGA 239

Db 880 TGAGCCACTACAAGCTGTCTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGGCGGAGGA 821

QY 240 GAGAAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCA 299

Db 820 GAGAAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCA 761

QY 300 ACGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCGACCGCCTGGCCGTGC 359

Db 760 ACGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCGACCGCCTGGCCGTGC 701

QY 360 CGCCCTTCGCCAGCGGGAGCGCTTCACCGCTTCAGGCCACCTATCCGTACCTGCAGC 419

Db 700 CGCCCTTCGCCAGCGGGAGCGCTTCACCGCTTCAGGCCACCTATCCGTACCTGCAGC 641

QY 420 ACGAGATCGACCTGCCGCCACCATCTCGCTGTTCAGACGGGAGGAGCCGCCACCTACC 479

Db 640 ACGAGATCGACCTGCCGCCACCATCTTCGCTGTTCGACGGGAGGAGCCGCCACCTACC 581

QY 480 ACGGCCCTGCACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAAACCGGAGT 539

Db 580 ACGGCCCTGCACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAAACCGGAGT 521

QY 540 CGGTGCGCGCACCCCCAAACAGAACCATCTTCGACAGTACCTGATGGATAGTCCAGGC 599

Db 520 CGGTGCGCGCACCCCCAAACAGAACCATCTTCGACAGTACCTGATGGATAGTCCAGGC 461

QY 600 TGGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCG 659

Db 460 TGGGCGGCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCG 401

QY 660 GCGGCGCATGGAGGGGCGCG-CCACCTACAGCGAGTTCGGCCACTACCCGGGG 718

Db 400 GCGGCGCATGGAGGGGCGCGCGCCACCTACAGCGAGTTCGGCCACTACCCGGGG 341

QY 719 TCCTCCTTCCAGCACCCAGCAGCAGTGGGCGCCCTCTCTGAGGGGACCCGGCTC 778

Db 340 TCCTCCTTCCAGCACCCAGCAGCAGTGGGCGCCCTCTCTGAGGGGACCCGGCTC 281

QY 779 CACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAA 838

Db 280 CACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAA 221

QY 839 CAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCCGGGCTGGGGTGCCTAGGTGAAA 898

Db 220 CAGAAAGGACACCTCTCTAGGGTCCCAGGGGGCCGGGCTGGGGTGCCTAGGTGAAA 161

QY 899 AGGCAGAACTCCCGCTTCTTAGAAGAGGAGTGAAGAGGCGGGGGCGCAGCAAC 958

Db 160 AGGCAGAACTCCCGCTTCTTAGAAGAGGAGTGAAGAGGCGGGGGCGCAGCAAC 101

QY 959 GCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTCTGG 1018

Db 100 GCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTCTGG 41

QY 1019 TCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1058

Db 40 TCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1

RESULT 6

AAA75151

ID AAA75151 standard; cDNA; 969 BP.

XX

AC AAA75151;

XX

DT 15-JAN-2001 (first entry)

XX

DE cDNA encoding a human TANGO 261 polypeptide.

XX

KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 6..764
FT /*tag= a
FT /product= "TANGO 261"
FT sig_peptide 6..89
FT /*tag= b
FT mat_peptide 90..764
FT /*tag= c
XX
PN WO200052022-A1.
XX
PD 08-SEP-2000.
XX
XX
PF 01-MAR-2000; 2000WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
PI WPI; 2000-579269/54.
XX P-PSDB; AAB18449.
PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -
XX
XX Claim 2; Fig 5; 175pp; English.
PS
XX
CC The present sequence encodes a human TANGO 261 polypeptide. The
CC specification also describes TANGO 266, TANGO 216, TANGO 262, and
CC TANGO 267. The TANGO polypeptides can be used to modulate cellular
CC proliferation, modulate cellular differentiation and/or modulate
CC cellular adhesion. The proteins can be used to treat any von Willebrand
CC factor-associated disorder, regulate extracellular matrix structuring,
CC cellular adhesion, and cell trafficking and/or migration, modulate
CC cellular interactions, modulate cell adhesion in proliferative
CC disorders, such as cancer, modulate the proliferation, differentiation,
CC and/or function of cells that appear in the bone marrow, and leukocytes,
CC treat bone marrow, blood and hematopoietic associated diseases and
CC disorders, atelectasis, pulmonary congestion or oedema, emphysema,
CC chronic bronchitis, bronchial asthma and bronchiectasis, intestinal
CC disorders, spleen associated diseases, modulate renal disorders, treat
CC cardiovascular disorders such as ischemic heart disease, modulate the
CC proliferation, differentiation, and/or function of bone and cartilage
CC cells and to treat bone and/or cartilage associated diseases or
CC disorder. They may also be used to treat disorders associated with the
CC ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic
CC disease, inflammations, bacterial and viral meningitis, Alzheimer's
CC Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,
CC brain cancers, hydrocephalus and encephalitis, and treat hepatic
CC disorders.
XX
SQ Sequence 969 BP; 211 A; 316 C; 294 G; 148 T; 0 other;

Query Match 90.7%; Score 961.8; DB 21; Length 969;
Best Local Similarity 99.8%; Pred. No. 1.3e-198;
Matches 963; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 7
AAA75163
ID AAA75163 standard; cDNA; 969 BP.
XX

QY 97 GGCAATGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGTGATGATGATGAT 156
DB 2 GGAGATGCGCGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGTGATGATGATGAT 61
QY 157 GGTGGTGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAG 216
DB 62 GGTGGTGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAG 121
QY 217 CCGGCACAGCAGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCTGTGGCC 276
DB 122 CCGGCACAGCAGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCTGTGGCC 181
QY 277 CTCGGAGAGACAGTGTTCAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCTCG 336
DB 182 CTCGGAGAGACAGTGTTCAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCTCG 241
QY 337 GCCACCGACCGCTGGCCGCTGCCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCA 396
DB 242 GCCACCGACCGCTGGCCGCTGCCGCCCTTCGCCAGCGGAGCGCTTCCACCGCTTCCA 301
QY 397 GCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGCTGTGAGA 456
DB 302 GCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGCTGTGAGA 361
QY 457 CGGGAGGAGCCCCACCTACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCA 516
DB 362 CGGGAGGAGCCCCACCTACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCA 421
QY 517 GCAGCTGGAACCTGAACCGGGAGTGGTGGCGGACCCCCCAACAGAACCATCTTCGACAG 576
DB 422 GCAGCTGGAACCTGAACCGGGAGTGGTGGCGGACCCCCCAACAGAACCATCTTCGACAG 481
QY 577 TGACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAGCAGTAACCTCGGCGAT 636
DB 482 TGACCTGATGATAGTGCAGGCTGGGCGGCCCTGCCCCCCCAGCAGTAACCTCGGCGAT 541
QY 637 CAGCGCCACGTGCTACGGCAGCGGGCGCGCATGGAGGGGCGCGCCACCTACAGCGCA 696
DB 542 CAGCGCCACGTGCTACGGCAGCGGGCGCGCATGGAGGGGCGCGCCACCTACAGCGCA 601
QY 697 GGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAGTGGGCGCGCTC 756
DB 602 GGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAGTGGGCGCGCTC 661
QY 757 CTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCAT 816
DB 662 CTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCAT 721
QY 817 CTGGAGCAAAAGAGAGATAAACAGAAAGGACACCTCTTAGGGTCCCCAGGGGGCGG 876
DB 722 CTGGAGCAAAAGAGAGATAAACAGAAAGGACACCTCTTAGGGTCCCCAGGGGGCGG 781
QY 877 GGCTGGGGCTGCGTAGTGAAAGGAGCAGAACACTCGCGCTTCTTAGAAGAGAGTGAGA 936
DB 782 GGCTGGGGCTGCGTAGTGAAAGGAGCAGAACACTCGCGCTTCTTAGAAGAGAGTGAGA 841
QY 937 GGAAGGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATA 996
DB 842 GGAAGGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATA 901
QY 997 AATATTACATGTGATGTCTGCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAAAA 1056
DB 902 AATATTACATGTGATGTCTGCTGAATGCACAAGCTTAAGAGAGCTTGCAAAAAAAA 961
QY 1057 AAAAA 1061
DB 962 AAAAA 966

AC AAA75163;
XX 15-JAN-2001 (first entry)
DT
XX cDNA clone encoding a human TANGO 261 polypeptide.
DE
XX
KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 6..764
FT /*tag= a
FT /product= "TANGO 261"
XX
PN WO200052022-A1.
XX
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
DR WPI; 2000-579269/54.
DR P-PSDB; AAB18461.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -
XX
PS Disclosure; Page -; 175pp; English.
XX
CC AAA75163-65 encode human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.
XX
SQ Sequence 969 BP; 211 A; 317 C; 293 G; 148 T; 0 other;

Query Match 90.5%; Score 960.2; DB 21; Length 969;

RESULT 8

Best Local Similarity 99.7%; Pred. No. 2.8e-198; Matches 962; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	97	GGCAATGGCGGAGCTGGAGTTGTTTTCAGATCATCATCTCGTGGTGGTATGATGATGAT	156
Db	2	GGAGATGGCGGACCTGGAGTTGTTTCAGATCATCATCTCGTGGTGGTATGATGATGAT	61
Qy	157	GGTGGTGGTATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATCAG	216
Db	62	GGTGGTGGTATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATCAG	121
Qy	217	CCGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCCTGTGGCC	276
Db	122	CCGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCCTGTGGCC	181
Qy	277	CTCGGAGACACAGTGTACGCAACGGAAATCCCAGAGCCCGCAGGTCTACGCCCGCCTCG	336
Db	182	CTCGGAGACACAGTGTACGCAACGGAAATCCCAGAGCCCGCAGGTCTACGCCCGCCTCG	241
Qy	337	GCCACCGACCGCTGGCCGTGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCA	396
Db	242	GCCACCGACCGCTGGCCGTGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCA	301
Qy	397	GCCACCTATCCGTACCTGCAGACAGAGATCGACCTGCCGCCACCATCTCGTGTGAGA	456
Db	302	GCCACCTATCCGTACCTGCAGACAGAGATCGACCTGCCGCCACCATCTCGTGTGAGA	361
Qy	457	CGGGAGGAGCCCCACCTTACAGGGCCCTTGACCCCTCCAGCTTCGGACCCCGAGCA	516
Db	362	CGGGAGGAGCCCCACCTTACAGGGCCCTTGACCCCTCCAGCTTCGGACCCCGAGCA	421
Qy	517	GCAGCTGGAACCTGAACCGGGAGTGGTGCAGCCACCCCAACAGAACCATCTTCGACAG	576
Db	422	GCAGCTGGAACCTGAACCGGGAGTGGTGCAGCCACCCCAACAGAACCATCTTCGACAG	481
Qy	577	TGACCTGATGGATAGTGCAGGGTGGGGGGCCCTGCCCCCAGCAGTAACCTCGGGCAT	636
Db	482	TGACCTGATGGATAGTGCAGGGTGGGGGGCCCTGCCCCCAGCAGTAACCTCGGGCAT	541
Qy	637	CAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCCGCCACCTACAGCGA	696
Db	542	CAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCCGCCACCTACAGCGA	601
Qy	697	GGTCATCGGCCACTACCGGGGCTCTCTTCCAGCACCAAGCAGAGCAGTGGCGCCCTC	756
Db	602	GGTCATCGGCCACTACCGGGGCTCTCTTCCAGCACCAAGCAGAGCAGTGGCGCCCTC	661
Qy	757	CTTGCTGGAGGGACCCGGCTCCACACACACACATCGCGCCCTTAGAGAGCGCAGCCAT	816
Db	662	CTTGCTGGAGGGACCCGGCTCCACACACACACATCGCGCCCTTAGAGAGCGCAGCCAT	721
Qy	817	CTGAGCAAAAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCCCAGGGGGCCG	876
Db	722	CTGAGCAAAAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCCCAGGGGGCCG	781
Qy	877	GGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGGCTTCTTAGAAGAGGAGTGAGA	936
Db	782	GGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGGCTTCTTAGAAGAGGAGTGAGA	841
Qy	937	GGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATA	996
Db	842	GGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATA	901
Qy	997	AATATTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1056
Db	902	AATATTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	961
Qy	1057	AAAAA 1061	
Db	962	AAAAA 966	

AAA75164
ID AAA75164 standard; cDNA; 969 BP.
XX
AC AAA75164;
XX
DT 15-JAN-2001 (first entry)
XX
DE cDNA clone encoding a human TANGO 261 polypeptide.
XX
KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 6..764
FT /*tag= a
FT /product= "TANGO 261"
XX
WO200052022-A1.
PN
PD 08-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US05226.
XX
PR 01-MAR-1999; 99US-0122458.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
XX
DR WPI; 2000-579269/54.
DR P-PSDB; AAB18462.
XX
PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -
XX
PS Disclosure; Page -; 175pp; English.
XX
CC AAA75163-65 encode human TANGO 261 proteins. The specification also
CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO
CC polypeptides can be used to modulate cellular proliferation, modulate
CC cellular differentiation and/or modulate cellular adhesion. The
CC proteins can be used to treat any von Willebrand factor-associated
CC disorder, regulate extracellular matrix structuring, cellular adhesion,
CC and cell trafficking and/or migration, modulate cellular interactions,
CC modulate cell adhesion in proliferative disorders, such as cancer,
CC modulate the proliferation, differentiation, and/or function of cells
CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood
CC and hematopoietic associated diseases and disorders, atelectasis,
CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial
CC asthma and bronchiectasis, intestinal disorders, spleen associated
CC diseases, modulate renal disorders, treat cardiovascular disorders such
CC as ischemic heart disease, modulate the proliferation, differentiation,
CC and/or function of bone and cartilage cells and to treat bone and/or
CC cartilage associated diseases or disorder. They may also be used to
CC treat disorders associated with the ovaries, and cerebral oedema,
CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,
CC bacterial and viral meningitis, Alzheimer's Disease, cerebral
CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,
CC hydrocephalus and encephalitis, and treat hepatic disorders.
CC note: the present sequence does not appear in the specification; it was
CC created using information provided.

SQ Sequence 969 BP; 210 A; 317 C; 294 G; 148 T; 0 other;
Query Match 90.5%; Score 960.2; DB 21; Length 969;
Best Local Similarity 99.7%; Pred. No. 2.8e-198;
Matches 962; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 97 GGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGTGGTATGATGGTGAT 156
Db 2 GGAGATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTCGTGGTGGTATGATGGTGAT 61
QY 157 GGTGGTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTCTTCATCAG 216
Db 62 GGTGGTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTCTTCATCAG 121
QY 217 CCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCCTGTGGCC 276
Db 122 CCGGCACAGCCAGGGCGGAGGAGAGACGATGCCCTGTCTCAGAAAGGATGCCCTGTGGCC 181
QY 277 CTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGACGGTCTACGCCCCGCCCTCG 336
Db 182 CTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGACGGTCTACGCCCCGCCCTCG 241
QY 337 GCCCACCGACCGCCTGGCCGTGCCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCA 396
Db 242 GCCCACCGACCGCCTGGCCGTGCCGCCCTTCCGCCAGCGGAGCGCTTCCACCGCTTCCA 301
QY 397 GCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTGAGA 456
Db 302 GCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTGAGA 361
QY 457 CGGGGAGGAGCCCCCACCTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCA 516
Db 362 CGGGGAGGAGCCCCCACCTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCA 421
QY 517 GCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCACCCCCAAAACAGAACCATCTTCGACAG 576
Db 422 GCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCACCCCCAAAACAGAACCATCTTCGACAG 481
QY 577 TGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTTGCCGCCACAGCAGTAACCTCGGGCAT 636
Db 482 TGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTTGCCGCCACAGCAGTAACCTCGGGCAT 541
QY 637 CAGGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGCCCGCCGCCACCTACAGCGA 696
Db 542 CAGGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGCCCGCCGCCACCTACAGCGA 601
QY 697 GGTCAATCGGCCACTACCCGGGTCTCTCTCCAGCACAGCAGCAGTGGGCCGCCCTC 756
Db 602 GGTCAATCGGCCACTACCCGGGTCTCTCTCCAGCACAGCAGCAGTGGGCCGCCCTC 661
QY 757 CTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCAT 816
Db 662 CTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCAT 721
QY 817 CTGAGCAAAAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGG 876
Db 722 CTGAGCAAAAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCGG 781
QY 877 GGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGAGTGAGA 936
Db 782 GGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGAGTGAGA 841
QY 937 GGAAGCGGGGGCGCAGCAACGCATCGTGTGGGCCCTCCCTCCACCTCCCTGTGTATA 996
Db 842 GGAAGCGGGGGCGCAGCAACGCATCGTGTGGGCCCTCCCTCCACCTCCCTGTGTATA 901
QY 997 AATATTTACATGTCTGTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAANA 1056
Db 902 AATATTTACATGTCTGTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAANA 961
QY 1057 AAAAA 1061
Db 962 AAAAA 966

RESULT 9

AAA75165
ID AAA75165 standard; cDNA; 969 BP.

XX
AC

XX
DT 15-JAN-2001 (first entry)

XX DE CDNA clone encoding a human TANGO 261 polypeptide.

XX
KW TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX
OS Homo sapiens.

XX	Key
FH	CDS
FT	
FT	Location/Qualifiers
FT	6..764
FT	/*tag= a
FT	/product= "TANGO 261"

XX
PN WO200052022-A1.

XX
PD 08-SEP-2000.

XX
PF 01-MAR-2000; 2000WO-US05226.

XX
PR 01-MAR-1999; 99US-0122458.

XX
PA (MILL-) MILLENNIUM PHARM INC.

XX Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

XX
DR WPI; 2000-579269/54.

DR P-PSDB; AAB18463.
XX

PT Novel human and murine secreted proteins designated TANGO 216, 261,
PT 262, 266 and 267 useful as modulating agents of cellular processes,
PT e.g. for treating cancer -

XX
PS Disclosure; Page -; 175pp; English.

AAA75163-65 encode human TANGO 261 proteins. The specification also describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO polypeptides can be used to modulate cellular proliferation, modulate cellular differentiation and/or modulate cellular adhesion. The proteins can be used to treat any von Willebrand factor-associated disorder, regulate extracellular matrix structuring, cellular adhesion, and cell trafficking and/or migration, modulate cellular interactions, modulate cell adhesion in proliferative disorders, such as cancer, modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, and leukocytes, treat bone marrow, blood and hematopoietic associated diseases and disorders, atelectasis, pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial asthma and bronchiectasis, intestinal disorders, spleen associated diseases, modulate renal disorders, treat cardiovascular disorders such as ischemic heart disease, modulate the proliferation, differentiation, and/or function of bone and cartilage cells and to treat bone and/or cartilage associated diseases or disorder. They may also be used to treat disorders associated with the ovaries, and cerebral oedema, hydrocephalus, brain herniations, iatrogenic disease, inflammations, bacterial and viral meningitis, Alzheimer's Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers, hydrocephalus and encephalitis, and treat hepatic disorders.

QY 202 ACGTGCTTTCATCAGCCGGCACAGCCAGGGGGAGGAGAGAGATGCCCTGTCTCAGA 261
Db 214 CCGCTCCTTCATCAGCCGACACAGCCAGGCCAGGAGAGAGATGGACTGTCTCGGA 273
QY 262 AGGATGCTGTGGCCCTCGGAGAGACAGTGTTCAGGCAACGGAATCCAGAGCCGAGGT 321
Db 274 AGGATGCTGTGGCCCTCAGAGAGTACGGTGTTCAGG--TGAATGCCGGAGCCACAGGT 330
QY 322 CTACGCCCGCCTCGGCCACCGACCGCCCTGGCCGTGGCCGCTTTCGCCCCAGCGGGAGCG 381
Db 331 CTATGCCCGCCTCGGCCACTGACCGACTCGTGTGCCCCCTTCATCCAGCGG----- 385
QY 382 CTTCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACTGCCGCCAC 441
Db 386 ----AGCCGATTCCAACCCACCTACCCCTACCTGCAGCACGAATTCCTTGCACCCAC 441
QY 442 CATCTCGTGTTCAGCGGGAGGAGGCCCCACCCCTACCGAGGCCCTGCACCCCTCCAGCT 501
Db 442 CATCTCACTGTCTGATGGGAGGAGGCCCCACCCCTACCGAGGCCCTGCACCCCTCCAGCT 501
QY 502 TCGGGACCCCGAGCAGCTGGAACCTGAACCGGGAGTTCGGTGGCGCACCCCAACAG 561
Db 502 ACGGGACCCCTGAGCAACAGCTGGAGCTGAACCGGAATCTGTGGCGCACCCCTTAACCG 561
QY 562 AACCATCTTCGACAGTGACCTGATGTAGTCCAGGCTGGCGGCCCTGCCCCCCAG 621
Db 562 GACCATCTTCGACAGTGACCTTATAGACAGCACCATGTCTGGGGGCCCTGTCCCCCAG 621
QY 622 CAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGCCGCC 681
Db 622 CAGTAACCTCGGGCATCAGCGCCACCTGCTACAGCAGCGGTGGCGGCATGGAGGGCCGCC 681
QY 682 GCCACCTACAGCGAGGTATCGGCCACTACCCGGGTCTCCTTCCAGCACACAGAG 741
Db 682 CCCCACCTACAGCGAGGTATTCGGCCACTACCTGGCTCCTCTCCAGCACCAAG 741
QY 742 CAGTGGCGCCCTCCTTGTGGAGGGGACCCGGCTCCACACACACATCGCGCCCT 801
Db 742 TAACGGSCCATCTCCTCTGCTAGAGGGACCCGGCTCCATCATCTGCACATTGCCCCACT 801
QY 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAAGAGGACACCCCTCTCTAGGG 861
Db 802 GGA-----GAACAGGAGAGGAGAAACAGAAAGTCAACCCCTCTAGGA 846
QY 862 TCCCGAGGGGGCGG 876
Db 847 GTGGGGGCGGGGCG 861

RESULT 14

AAA75152

ID AAA75152 standard; cDNA; 1713 BP.

XX AC AAA75152;

XX AC AAA75152;

DT 15-JAN-2001 (first entry)

XX DE cDNA encoding a murine TANGO 261 polypeptide.

DE cDNA encoding a murine TANGO 261 polypeptide.

XX TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
KW cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; bronchial asthma; bronchiectasis;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.
OS Mus sp.
XX Key
FH Location/Qualifiers

CDS 2..655
/*tag= a
/product= "TANGO 261"
WO200052022-A1.
08-SEP-2000.
01-MAR-2000; 2000WO-US05226.
01-MAR-1999; 99US-0122458.
(MILL-) MILLENNIUM PHARM INC.
Barnes TM, Holtzman DA, Sharp JD, Fraser CC;
WPI; 2000-579269/54.
P-PSDB; AAB18450.
Novel human and murine secreted proteins designated TANGO 216, 261,
262, 266 and 267 useful as modulating agents of cellular processes,
e.g. for treating cancer -
Claim 2; Fig 6A-B; 175pp; English.
The present sequence encodes a murine TANGO 261 polypeptide. The
specification also describes TANGO 266, TANGO 216, TANGO 262, and
TANGO 267. The TANGO polypeptides can be used to modulate cellular
proliferation, modulate cellular differentiation and/or modulate
cellular adhesion. The proteins can be used to treat any von Willebrand
factor-associated disorder, regulate extracellular matrix structuring,
cellular adhesion, and cell trafficking and/or migration, modulate
cellular interactions, modulate cell adhesion in proliferative
disorders, such as cancer, modulate the proliferation, differentiation,
and/or function of cells that appear in the bone marrow, and leukocytes,
treat bone marrow, blood and hematopoietic associated diseases and
disorders, atelectasis, pulmonary congestion or oedema, emphysema,
chronic bronchitis, bronchial asthma and bronchiectasis, intestinal
disorders, spleen associated diseases, modulate renal disorders, treat
cardiovascular disorders such as ischemic heart disease, modulate the
proliferation, differentiation, and/or function of bone and cartilage
cells and to treat bone and/or cartilage associated diseases or
disorder. They may also be used to treat disorders associated with the
ovaries, cerebral oedema, hydrocephalus, brain herniations, iatrogenic
disease, inflammations, bacterial and viral meningitis, Alzheimer's
Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis,
brain cancers, hydrocephalus and encephalitis, and treat hepatic
disorders.
XX
SQ Sequence 1713 BP; 506 A; 438 C; 400 G; 369 T; 0 other;

Query Match 49.1%; Score 521.4; DB 21; Length 1713;
Best Local Similarity 78.5%; Pred. No. 2.5e-103;
Matches 704; Conservative 0; Mismatches 151; Indels 42; Gaps 5;

Qy 178 GCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCAGCCAGGGCGGAG 237
Db 1 GCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGGCAGCCAGGGCGGAG 60
Qy 238 GAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGAGACAGTGTACAG 297
Db 61 GAGAGACGATGGACTGTCTCGGAAGGATGCCTGTGGCCCTCAGAGAGTACGGTGTACAG 120
Qy 298 CAACGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCGACCGCTGGCCGT 357
Db 121 ---TGAATGCCGGAGCCACAGGTCTATGCCCGCCCTCGGCCACTGACCGACTCGCTGT 177
Qy 358 GCGGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCA 417
Db 178 GCGGCCCTTCATCCAGCGG-----AGCCGATTCACACCCACCTACCCCTACCTGCA 228
Qy 418 GCACGAGATCGACCTGCCGCCACCATCTCGTGTCTCAGCGGGGAGGAGCCCCCCTA 477

Db 229 GCACGAAATTGCCCTGCCACCCACCATCTCACTGTCTGATGGGAGGAGCCCCACCCCTA 288

QY 478 CCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCGGGA 537
|||||

Db 289 CCAGGGCCCTGCACCCCTCCAGCTACGGGACCTTGAGCAACAGCTGGAGCTGAACCGGGA 348
|||||

QY 538 GTCGTGCGCGCACCCCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCAG 597
|||||

Db 349 ATCTGTGCGCGCACCCCTTAACCGACCATCTTCGACAGTGACCTTATAGACAGCACCAT 408
|||||

QY 598 GCTGGGCGGCCCTGCCCCCCACGACGATACTCGGGCATACGGCCACGTGTACGGCAG 657
|||||

Db 409 GCTGGGCGGCCCTGTCCCCCAGCAGTAACCTCGGGCATACGGCCACCTGTCTACAGCAG 468
|||||

QY 658 CGGGGGCGCATGGAGGGCGCGCGCCACCTACAGCGAGGTATCGGCCACTACCCGGG 717
|||||

Db 469 CGGTGGGCGCATGGAGGGCGCGCGCCACCTACAGCGAGGTATCGGCCACTACCCCTGG 528
|||||

QY 718 GTCCTCCTCCAGCACCGACGAGCAGTGGCGCGCCCTCCTTGCTGGAGGGACCCCGCT 777
|||||

Db 529 CTCCTCCTCCAGCACCGACCAAGTAACGGCCATCTCTCCCTGCTAGAGGGACCCCGCT 588
|||||

QY 778 CCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAA 837
|||||

Db 589 CCATCACTCGCACATTGCCCCACTGGA-----GAACAAGGAGAAGGAGAA 633
|||||

QY 838 ACAGAAAGGACACCCCTCTAGGTCCCCAGGGGGCGCGGCTGGGGCTGCGTAGGTGAA 897
|||||

Db 634 ACAGAAAGGTACCCCTCTAGAGTGGGGCGCGGGCGCTGTAGGCAAAACCGCAAAA 693
|||||

QY 898 AAGGCAG-----AACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCG 944
|||||

Db 694 AAAAAAGAAAGGAAAAAAMAACTCCGCACTTCTTAAGAGAGAAGAGAGGAAAGTCA 753
|||||

QY 945 GGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTTA 1004
|||||

Db 754 GGGGACACA--CAGGCTGAGTGGCGGTGGTAGTTCTCCTGTCTGTGTATAAATATTTA 811
|||||

QY 1005 CATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAMAAAAA 1061
|||||

Db 812 CATGTTCTGTGTGGTCTGAATGCAGAAGCTCAAAAAGCTTGCAAAAAAAGAAAAAGA 868
|||||

RESULT 15
AAA75166
ID AAA75166 standard; cDNA; 1713 BP.

AC AAA75166;

XX 15-JAN-2001 (first entry)

DT cDNA clone encoding a murine TANGO 261 polypeptide.

DE TANGO 266; TANGO 216;; TANGO 261; TANGO 262; TANGO 267;
XX cellular proliferation; cellular differentiation; cellular adhesion;
KW von Willebrand factor-associated disorder; cell trafficking; cancer;
KW hematopoietic associated disease; atelectasis; pulmonary congestion;
KW oedema; emphysema; chronic bronchitis; bronchial asthma; bronchiectasis;
KW intestinal disorder; spleen associated disease; renal disorder;
KW cardiovascular disorder; ischemic heart disease; hydrocephalus;
KW brain herniation; iatrogenic disease; inflammation; meningitis;
KW Alzheimer's Disease; cerebral toxoplasmosis; Parkinson's disease;
KW multiple sclerosis; hydrocephalus; encephalitis; hepatic disorder; ss.

XX Mus sp.

OS Location/Qualifiers

XX Key 2..655

FT CDS /*tag= a

FT /product= "TANGO 261"

XX WO200052022-A1.

PN

XX

PD 08-SEP-2000.

XX 01-MAR-2000; 2000WO-US05226.

PF 01-MAR-1999; 99US-0122458.

XX (MILL-) MILLENNIUM PHARM INC.

PA Barnes TM, Holtzman DA, Sharp JD, Fraser CC;

PI WPI; 2000-579269/54.

XX P-PSDB; AAB18464.

DR Novel human and murine secreted proteins designated TANGO 216, 261,

PT 262, 266 and 267 useful as modulating agents of cellular processes,

PT e.g. for treating cancer -

XX Disclosure; Page -: 175pp; English.

PS

XX AAA75166-68 encode murine TANGO 261 proteins. The specification also

CC describes TANGO 266, TANGO 216, TANGO 262, and TANGO 267. The TANGO

CC polypeptides can be used to modulate cellular proliferation, modulate

CC cellular differentiation and/or modulate cellular adhesion. The

CC proteins can be used to treat any von Willebrand factor-associated

CC disorder, regulate extracellular matrix structuring, cellular adhesion,

CC and cell trafficking and/or migration, modulate cellular interactions,

CC modulate cell adhesion in proliferative disorders, such as cancer,

CC modulate the proliferation, differentiation, and/or function of cells

CC that appear in the bone marrow, and leukocytes, treat bone marrow, blood

CC and hematopoietic associated diseases and disorders, atelectasis,

CC pulmonary congestion or oedema, emphysema, chronic bronchitis, bronchial

CC asthma and bronchiectasis, intestinal disorders, spleen associated

CC diseases, modulate renal disorders, treat cardiovascular disorders such

CC as ischemic heart disease, modulate the proliferation, differentiation,

CC and/or function of bone and cartilage cells and to treat bone and/or

CC cartilage associated diseases or disorder. They may also be used to

CC treat disorders associated with the ovaries, and cerebral oedema,

CC hydrocephalus, brain herniations, iatrogenic disease, inflammations,

CC bacterial and viral meningitis, Alzheimer's Disease, cerebral

CC toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers,

CC hydrocephalus and encephalitis, and treat hepatic disorders.

CC note: the present sequence does not appear in the specification; it was

CC created using information provided.

XX

SQ Sequence 1713 BP; 505 A; 439 C; 400 G; 369 T; 0 other;

Query Match 49.0%; Score 519.8; DB 21; Length 1713;
Best Local Similarity 78.4%; Pred. No. 5.5e-103;
Matches 703; Conservative 0; Mismatches 152; Indels 42; Gaps 5;

QY 178 GCTGAGCCACTACAAGCTGTCTGCACGGTCCCTTCATCAGCCGACAGCCGGCGGAG 237
|||||

Db 1 GCTGAGCCACTACAAGCTGTCTGAGCCCGCTCTTCATCAGCCGACAGCCAGGCGGAG 60
|||||

QY 238 GAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTGAG 297
|||||

Db 61 GAGAGACGATGGACTGTCTCTGGACGGATGCCTCTGGCCCTCAGAGAGTACGGTGTGAG 120
|||||

QY 298 CAACGGAATCCAGAGCCGAGGTTCTACGCCCGCTCGGCCACCGACCGCTGGCCGT 357
|||||

Db 121 ---TGAATGCCGGAGCCACAGGTCTATGCCCGCTCGGCCCTGACTGACCGACTCGCTGT 177
|||||

QY 358 GCCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCAGACCCACCTATCCGTACCTGCA 417
|||||

Db 178 GCCCCCTTCATCCAGCGG-----AGCCGATTCCCAACCCACCTACCCCTACCTGCA 228
|||||

QY 418 GCACGAGATCGACCTGCCGCCACCACCATCTCGCTGTCTAGACGGGAGGAGCCCCCACCCTA 477
|||||

Db 229 GCACGAAATTGCCCTGCCACCCACCATCTCACTGTCTGTATGGGAGGAGCCCCCACCCTA 288
|||||

QY 478 CCAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGA 537
|||||

Db 289 CCAGGGCCCTGCACCCCTCCAGCTACGGGACCCCTGAGCAACAGCTGGAGCTGAACCGGGA 348
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 21:21:44 ; Search time 4130 Seconds
(without alignments)
10509.713 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tcctccttggttcggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues 5777422
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1034.4	97.5	1141	9	AF224278	AF224278 Homo sapi
2	1032.2	97.3	1818	9	AY128643	AY128643 Homo sapi
3	957.6	90.3	1061	9	BC015918	BC015918 Homo sapi
4	956	90.1	4839	9	AF305616	AF305616 Homo sapi
5	800.6	75.5	1321	6	AX392417	AX392417 Sequence
6	752.6	70.9	861	6	AX392419	AX392419 Sequence
7	748.2	70.5	61505	9	AF305426	AF305426 Homo sapi
8	728.8	68.7	1583	6	AX593655	AL035541 Human DNA
9	728.8	68.7	1583	6	AX593655	AX593655 Sequence
10	644.2	60.7	693	6	AX392430	AX392430 Sequence
11	584	55.0	1379	10	BC036995	BC036995 Mus muscu
12	523.8	49.4	878	6	AX392428	AX392428 Sequence
13	433.4	40.8	156698	10	AL837509	AL837509 Mouse DNA
14	433.4	40.8	175754	2	AC110189	AC110189 Mus muscu
15	433.4	40.8	176821	2	AL837520	AL837520 Mus muscu
16	423.4	39.9	651	10	AF220208	AF220208 Mus muscu
17	421.4	39.7	249554	2	AC139417	AC139417 Rattus no
18	421.4	39.7	258632	2	AC111878	AC111878 Rattus no
19	399.8	37.7	408	6	AX071267	AX071267 Sequence
20	399	37.6	673	6	AX525744	AX525744 Sequence
21	365.6	34.5	812	6	AX011709	AX011709 Sequence
22	350.8	33.1	2170	6	AX713513	AX713513 Sequence
23	350.8	33.1	2170	9	AK055028	AK055028 Homo sapi
24	350	33.0	8093	6	AX392432	AX392432 Sequence
25	350	33.0	8093	9	AF009426	AF009426 Homo sapi
26	349	32.9	921	6	AR233384	AR233384 Sequence
27	349	32.9	8065	6	AR233383	AR233383 Sequence
28	349	32.9	8494	9	AF009424	AF009424 Homo sapi
29	282	26.6	8039	9	AF009427	AF009427 Homo sapi
30	281	26.5	867	6	AR233385	AR233385 Sequence
31	281	26.5	1573	9	BC030199	BC030199 Homo sapi
32	281	26.5	8440	9	AF009425	AF009425 Homo sapi
33	250.2	23.6	172692	2	AP001013	AP001013 Homo sapi
34	250.2	23.6	173709	9	AP001010	AP001010 Homo sapi
35	250.2	23.6	181083	2	AP001268	AP001268 Homo sapi
36	249.2	23.5	475	6	AX392431	AX392431 Sequence
37	240.8	22.7	766	10	BC022716	BC022716 Mus muscu
38	232.6	21.9	240453	2	AC097603	AC097603 Rattus no
39	232.6	21.9	280575	2	AC117364	AC117364 Rattus no
40	231	21.8	183681	2	AC111069	AC111069 Mus muscu
41	223.8	21.1	167489	2	BX005347	BX005347 Danio rer
42	206.4	19.5	176458	2	AC134911	AC134911 Mus muscu
43	182.4	17.2	155348	5	AL928820	AL928820 Zebrafish
44	158.6	14.9	150224	9	HSJ1059L7	AL121913 Human DNA
45	103.6	9.8	2570	9	AK056098	AK056098 Homo sapi

ALIGNMENTS

RESULT 1
AF224278
LOCUS AF224278 1141 bp mRNA linear PRI 18-JUL-2000
DEFINITION Homo sapiens PMEPA1 protein (PMEPA1) mRNA, complete cds.
ACCESSION AF224278
VERSION AF224278.1 GI:9255808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1141)
AUTHORS Xu,L.L., Shanmugam,N., Segawa,T., Sesterhenn,I.A., McLeod,D.G.,
Moul,J.W. and Srivastava,S.
TITLE A novel androgen-regulated gene, PMEPA1, located on chromosome

FEATURES		Location/Qualifiers	
source		1..1818	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/chromosome="20"	
		/map="20q13.31-13.33"	
		145..858	
CDS		/codon_start=1	
		/product="PMEPA1 variant A protein"	
		/protein_id="AAW89277.1"	
		/db_xref="GI:22121999"	
		/translation="MMVMVVVITCLLSHYKLSARSFISRHSQRRRREDALSSEGCLWP SESTVSGNGIPEQVYVAPRPTRDLAVPPFAQERFHRFQPTYPYLQHEIDLPTISL SDGEPPPYQGQPCITQLRDPEQOLELNRESVRAPNRTIFDSDLMDSARIGGPCPPSS NSGISATCYGSGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAP LESAAIWSKEKDQKGHPL"	
BASE COUNT	459 a	467 c	494 g 398 t
ORIGIN			
Query Match 97.3%; Score 1032.2; DB 9; Length 1818;			
Best Local Similarity 99.1%; Pred. No. 2.3e-184;			
Matches 1051; Conservative 0; Mismatches 3; Indels 7; Gaps 1;			
QY	1	TCCTCCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC	60
Db	7	TCCTCCTTGGGTTCCGGTGAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC	66
QY	61	TGGAAGAACTGAAGCGGACGGTCTCTCGAAACAGGCAATGGCGGAGCTGGAGTTGT	120
Db	67	TGGAGAACTGAAGCGGACAGTCTCTCGAAACCA-----GCGGAGCTGGAGTTGT	119
QY	121	TCAGATCATCATCGTGGTGTGATGATGGTGTGATGGTGTGATCACGTGCCTGCT	180
Db	120	TCAGATCATCATCGTGGTGTGATGATGGTGTGATGGTGTGATCACGTGCCTGCT	179
QY	181	GAGCCACTACAAGCTGTCTGCACGGTCCTTCATCAGCCGCGCACAGCCAGGGCGGAGGAG	240
Db	180	GAGCCACTACAAGCTGTCTGCACGGTCCTTCATCAGCCGCGCACAGCCAGGGCGGAGGAG	239
QY	241	AGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAA	300
Db	240	AGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAA	299
QY	301	CGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCGACCGCCTGGCCGTGCC	360
Db	300	CGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCCACCGACCGCCTGGCCGTGCC	359
QY	361	GCCCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCA	420
Db	360	GCCCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCA	419
QY	421	CGAGATCGACCTGCCGCCACCATCTCGTGTTCAGACGGGAGGAGCCCCACCCCTACCA	480
Db	420	CGAGATCGACCTGCCGCCACCATCTCGTGTTCAGACGGGAGGAGCCCCACCCCTACCA	479
QY	481	GGGCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGTGGAACTGAACCGGGAGTC	540
Db	480	GGGCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGTGGAACTGAACCGGGAGTC	539
QY	541	GGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGATGATGCCAGGCT	600
Db	540	GGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGATGATGCCAGGCT	599
QY	601	GGGGGCCCTTGCCTCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGG	660
Db	600	GGGGGCCCTTGCCTCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACGGCAGCGG	659
QY	661	CGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTATCGGCCACTACCCGGGGTC	720
Db	660	CGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTATCGGCCACTACCCGGGGTC	719
QY	721	CTCCTTCAGCACACAGCAGAGCAGTGGGGCCGCCCTCCTTGTGTGGAGGGGACCCCGGCTCCA	780

Db	720	CTCCTTCCAGCACCCAGAGCAGTGGGCCCTCTCTGCTGGAGGGACCCGGCTCCA	779
QY	781	CCACACACATCGGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAGATAAACA	840
Db	780	CCACACACATCGGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAGATAAACA	839
QY	841	GAAGGACACCCCTCTTAGGGTCCCAGGGGGCCGGCTGGGCTGCGTAGGTGAAAAG	900
Db	840	GAAGGACACCCCTCTTAGGGTCCCAGGGGGCCGGCTGCGTAGGTGAAAAG	899
QY	901	GCAGAACACTCCGGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGC	960
Db	900	GCAGAACACTCCGGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGC	959
QY	961	ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTC	1020
Db	960	ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTC	1019
QY	1021	TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1061
Db	1020	TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1060
RESULT 3			
BC015918			
LOCUS			
DEFINITION Homo sapiens, clone MGC:20374 IMAGE:4559576, mRNA, complete cds.			
ACCESSION BC015918			
VERSION BC015918.1 GI:16198474			
KEYWORDS MGC.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 (bases 1 to 1061)			
AUTHORS Strausberg,R.			
TITLE Direct Submission			
JOURNAL Submitted (15-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT Contact: MGC help desk			
Email: cgapbs-r@mail.nih.gov			
Tissue Procurement: DCTD/DTP			
cDNA Library Preparation: Rubin Laboratory			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Genome Sequence Centre,			
BC Cancer Agency, Vancouver, BC, Canada			
info@bcgsc.bc.ca			
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			
Series: IRAL Plate: 29 Row: e Column: 5			
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.			
FEATURES			
source			
1..1061			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="MGC:20374 IMAGE:4559576"			
/tissue_type="Kidney, renal cell adenocarcinoma"			


```
|||||
Db 1099 GCCACCTACAGCGAGTTCATCGGCCACTACCGGGTCTCTCTCCAGCACGAG 1158
|||||
Qy 742 CAGTGGCGCCCTCTCTTGTGTGAGGGACCCGGCTCCACACACACATCGCGCCCT 801
|||||
Db 1159 CAGTGGCGCCCTCTCTTGTGTGAGGGACCCGGCTCCACACACACATCGCGCCCT 1218
|||||
Qy 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCCCTCTCTAGGG 861
|||||
Db 1219 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCCCTCTCTAGGG 1278
|||||
Qy 862 TCCCCAGGGGCGCGGCTGGGCTGCGTAGGTGAAAAGGCAG 904
|||||
Db 1279 TCCCCAGGGGCGCGGCTGGGCTGCGTAGGTGAAAAGGCAG 1321

RESULT 6
AX392419
LOCUS AX392419 861 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 3 from Patent WO0216416.
ACCESSION AX392419
VERSION AX392419.1 GI:19700734
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and
Turi,T.G.
TITLE Diagnosis and treatment of cardiovascular conditions
JOURNAL Patent: WO 0216416-A 3 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source
1. .861
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .>861
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD29006.1"
/db_xref="GI:19700735"
/translation="MHRLMGVNSTAAAAAGQPNVSTCNKRSFLQSMETLEFVQI
IIIVVMVMVVITCLSHYKLSARSFIRHSQRRRDALSSGCLWPSESTVSGN
GIPEPQVYAPRPTDLRVPFAQRFRFQTPYVQLQHEIDLPTISLSDGEEPPP
YQGPCTQLRDPQQLELNRESVRAPPNRTIFDSDLMSARLGGPCPPSNSGISATC
YSGGRMEGPPPTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIWS
KEKDKQKGHPL"
BASE COUNT 172 a 307 c 253 g 129 t
ORIGIN
Query Match 70.9%; Score 752.6; DB 6; Length 861;
Best Local Similarity 98.2%; Pred. No. le-131;
Matches 761; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 82 TCTCCTGCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 141
|||
Db 87 TTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 146
|||||
Qy 142 GGTGATGATGGTGGTGGTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGC 201
|||||
Db 147 GGTGATGATGGTGGTGGTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGC 206
|||||
Qy 202 ACGGTCCTTCATCAGCCGGCACAGCGGGGCGGAGGAGAGAGATGCCCTGTCTCAGA 261
|||||
Db 207 ACGGTCCTTCATCAGCCGGCACAGCGGGGCGGAGGAGAGAGATGCCCTGTCTCAGA 266
|||||
Qy 262 AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTGAGCAACGGGAATCCAGAGCCGCGAGT 321
|||||
Db 267 AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTGAGCAACGGGAATCCAGAGCCGCGAGT 326
|||||
Qy 322 CTACGCGCCGCTCGGCCCCACCGACCGCTGGCCGCTGCGCCCTTCGCCCGGAGCG 381
|||||
```

```
|||||
Db 327 CTAGGCCCCGCTCGGCCACCGACCGCCTGGCGTGC CGCCCTTCGCCCGAGCGGAGCG 386
|||||
Qy 382 CTTCCACCGCTTCAGCCCCACCTATCCGTACCTTCGACGACGAGATCGACCTGCGGCCAC 441
|||||
Db 387 CTTCCACCGCTTCAGCCCCACCTATCCGTACCTTCGACGACGAGATCGACCTGCCACCCAC 446
|||||
Qy 442 CATCTCGTGTGACGCGGGGAGGAGCCCCCACCCTACGAGGGCCCTGCACCTCCAGCT 501
|||||
Db 447 CATCTCGTGTGACGCGGGGAGGAGCCCCCACCCTACGAGGGCCCTGCACCTCCAGCT 506
|||||
Qy 502 TCGGACCCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGGCGGCACCCCAACACAG 561
|||||
Db 507 TCGGACCCCCGAGCAGCTGGAACCTGAACCGGAGTGGTGGCGGCACCCCAACACAG 566
|||||
Qy 562 AACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCGCCAG 621
|||||
Db 567 AACCATCTTCGACAGTACCTGATGGATAGTCCAGGCTGGGCGGCCCTGCCCGCCAG 626
|||||
Qy 622 CAGTAACTCGGCGATCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCC 681
|||||
Db 627 CAGTAACTCGGCGATCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCC 686
|||||
Qy 682 GCCCACCTACAGCGAGGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACGAGAG 741
|||||
Db 687 GCCCACCTACAGCGAGGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACGAGAG 746
|||||
Qy 742 CAGTGGGCGCCCTCTCTTGTGTGAGGGGACCCGGTCCACACACACATCGGCGCCCT 801
|||||
Db 747 CAGTGGGCGCCCTCTCTTGTGTGAGGGGACCCGGTCCACACACACATCGGCGCCCT 806
|||||
Qy 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCCCTCTC 856
|||||
Db 807 AGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCCCTCTC 861
|||||

RESULT 7
AF305426 61505 bp DNA linear PRI 12-OCT-2001
LOCUS AF305426 Homo sapiens solid tumor-associated 1 protein (STAG1/PMEP1) gene,
complete cds.
ACCESSION AF305426
VERSION AF305426.1 GI:15824468
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 61505)
AUTHORS Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
TITLE Identification and characterization of a novel gene, STAG1,
up-regulated in renal cell carcinoma and other solid tumours
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 61505)
AUTHORS Rae,F.K., Hooper,J.D., Nicol,D.L. and Clements,J.A.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2000) Centre for Molecular Biotechnology,
Queensland University of Technology, 2 George St, Brisbane, QLD
4001, Australia
FEATURES
source
1. .61505
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2-q13.33"
1. .61505
/gene="STAG1/PMEP1"
join(1. .429,50206. .50360,56817. .56870,57305. .61505)
/gene="STAG1/PMEP1"
/product="solid tumor-associated 1 protein"
join(321. .429,50206. .50360,56817. .56870,57305. .57850)
/gene="STAG1/PMEP1"
CDS
```

/codon_start=1
/product="solid tumor-associated 1 protein"
/protein_id="AA109357.1"
/db_xref="GI:15824469"
/translation="MHRLMGVNSTAAAAAGQPNVSCCTCNCKRSLFQSMETLEFVQI
IIIVVMVMVWVITCLLSHYKLSARSFISHSQGRREDLSSEGCLWPSESTVSGN
GIPEQVYAPRPTDLAVPPFAQRERFRFQTPYLOHEIDLPTTISLSDGEEPTP
YQGPCTQLRDPQQLNRESVRAPNRTIFDSDLMSARLGGPCPPSNGISATC
YSGGRMEGPPPTYSEVIGHYPGSSFQHQQSSPPSLLEGTRLHHTHIAPLESAAIWS
KEKDQKQGHPL"

BASE COUNT 13329 a 16074 c 17430 g 14672 t
ORIGIN

Query Match	70.5%	Score 748.2;	DB 9;	Length 61505;
Best Local Similarity	98.9%	Pred. No. 5.3e-131;		
Matches 753;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
Qy	301	CGGAATCCCGAGCGCGAGGCTACGCCCGCCCTCGGCCACCGACCGCCTGGCCGTGCC	360	
Db	57292	CTGCTTCTCCAGCCCGCAGGCTACGCCCGCCCTCGGCCACCGACCGCCTGGCCGTGCC	57351	
Qy	361	GCCCTTCGCCCGAGCGGGAGCGCTTCCACCGCTTCCAGCCCGACCTATCCGTACCTGCAGCA	420	
Db	57352	GCCCTTCGCCCGAGCGGGAGCGCTTCCACCGCTTCCAGCCCGACCTATCCGTACCTGCAGCA	57411	
Qy	421	CGAGATCGACCTGCCGCCCCACCATCTCGCTGTGCAGACGGGAGGAGCCCCACCCCTACCA	480	
Db	57412	CGAGATCGACCTGCCACCCACCATCTCGCTGTGCAGACGGGAGGAGCCCCACCCCTACCA	57471	
Qy	481	GGSCCCTGCACCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTC	540	
Db	57472	GGSCCCTGCACCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTC	57531	
Qy	541	GGTCGCGCACCCCCAAACAGAACCATCTTCGACAGTGACCTGATGGATAGTCCAGGCT	600	
Db	57532	GGTCGCGCACCCCCAAACAGAACCATCTTCGACAGTGACCTGATGGATAGTCCAGGCT	57591	
Qy	601	GGCGGCCCTTGCCCCCAGCAGCACTCGGGCCCGCTTCTGCTGGAGGGACCCCGCTCCA	780	
Db	57592	GGCGGCCCTTGCCCCCAGCAGCACTCGGGCCCGCTTCTGCTGGAGGGACCCCGCTCCA	57771	
Qy	661	CGGCGCATGGAGGGCGCGCCGACCATCTACAGCGAGTGCTATCGGCCACTACCCGGGTC	720	
Db	57652	CGGCGCATGGAGGGCGCGCCGACCATCTACAGCGAGTGCTATCGGCCACTACCCGGGTC	57711	
Qy	721	CTCCTTCCAGCACCAGCAGCACTGGGCCCGCCCTTCTGCTGGAGGGACCCCGCTCCA	780	
Db	57712	CTCCTTCCAGCACCAGCAGCACTGGGCCCGCCCTTCTGCTGGAGGGACCCCGCTCCA	57771	
Qy	781	CCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA	840	
Db	57772	CCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACA	57831	
Qy	841	GAAAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAG	900	
Db	57832	GAAAGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAG	57891	
Qy	901	GCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGC	960	
Db	57892	GCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGC	57951	
Qy	961	ATCGTGTGGCCCTCCCTCCCTCCCTGTGTATAAATATTACATGTGATGTCGTGTC	1020	
Db	57952	ATCGTGTGGCCCTCCCTCCCTCCCTGTGTATAAATATTACATGTGATGTCGTGTC	58011	
Qy	1021	TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1061	
Db	58012	TGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	58052	

RESULT 8
HS718J7/c
LOCUS HS718J7 130435 bp DNA linear PRI 24-FEB-2001

DEFINITION

Human DNA sequence from clone RP4-718J7 on chromosome 20q13.31-13.33 Contains the PKC1 gene for soluble phosphoenolpyruvate carboxykinase 1, part of a novel gene similar to mouse DLM-1 (tumour stroma and activated macrophage protein), the 3' end of the TMEPAI gene encoding an androgen induced lb transmembrane protein (PMEPAI), two putative novel genes, a CpG island, ESTs, STSS and GSSs, complete sequence.

ACCESSION

AL035541

VERSION

AL035541.15 GI:11546043

KEYWORDS

HTG; CpG island; DLM-1; macrophage protein; PKC1; phosphoenolpyruvate carboxykinase; PMEPAI; TMEPAI; transmembrane protein.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130435)

REFERENCE

AUTHORS

Sehra,H.

TITLE

Direct Submission

JOURNAL

Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

requests: clonerequest@sanger.ac.uk
On Dec 5, 2000 this sequence version replaced gi:10198628.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20

This sequence is the entire insert of clone RP4-718J7 The true left end of clone RP5-1007E6 is at 71437 in this sequence. The true right end of clone RP4-579F20 is at 43945 in this sequence.

RP4-718J7 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES

source

1. .130435
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="q13.31-13.33"
/clone="RP4-718J7"
/clone_lib="RPCI-4"

repeat_region

319..478
/note="5 copies 32 mer 67% conserved"

repeat_region

370..463
/note="2 copies 47 mer 87% conserved"

repeat_region

427..626
/note="4 copies 50 mer 71% conserved"

repeat_region

707..812
/note="53 copies 2 mer cc 61% conserved"

repeat_region

967..2752
/note="893 copies 2 mer gg 54% conserved"

repeat_region

982..2730
/note="33 copies 53 mer 54% conserved"


```
repeat_region 1177. .1820
/note="4 copies 161 mer 64% conserved"
repeat_region 1204. .1371
/note="3 copies 56 mer 75% conserved"
repeat_region 1358. .2757
/note="28 copies 50 mer 54% conserved"
repeat_region 1675. .2718
/note="18 copies 58 mer 55% conserved"
repeat_region 1843. .2346
/note="9 copies 56 mer 64% conserved"
repeat_region 1852. .2079
/note="4 copies 57 mer 86% conserved"
repeat_region 2028. .2632
/note="11 copies 55 mer 60% conserved"
repeat_region 2087. .2257
/note="3 copies 57 mer 83% conserved"
repeat_region 2228. .2497
/note="5 copies 54 mer 75% conserved"
repeat_region 2583. .2716
/note="2 copies 67 mer 82% conserved"
repeat_region 3378. .3426
/note="L1MB4 repeat: matches 6088. .6136 of consensus"
repeat_region 3799. .3896
/note="Charliel repeat: matches 681. .781 of consensus"
repeat_region 5331. .5793
/note="MLT1B repeat: matches 14. .466 of consensus"
repeat_region 5797. .5988
/note="6 copies 32 mer 86% conserved"
repeat_region 5855. .5978
/note="31 copies 4 mer gcac 61% conserved"
repeat_region 6474. .6591
/note="MIR repeat: matches 91. .218 of consensus"
repeat_region 6592. .6723
/note="FLAM_C repeat: matches 1. .132 of consensus"
repeat_region 6952. .7021
/note="L2 repeat: matches 2637. .2705 of consensus"
repeat_region 7358. .7671
/note="AluSx repeat: matches 1. .312 of consensus"
repeat_region 8521. .8554
/note="Alu repeat: matches 1. .34 of consensus"
repeat_region 8863. .9217
/note="MLT1D repeat: matches 105. .505 of consensus"
repeat_region 9228. .9746
/note="L2 repeat: matches 1459. .2028 of consensus"
repeat_region 9880. .10089
/note="L2 repeat: matches 2292. .2511 of consensus"
repeat_region 10094. .10206
/note="AluJb repeat: matches 188. .300 of consensus"
repeat_region 10263. .10573
/note="AluJb repeat: matches 1. .308 of consensus"
repeat_region 10658. .10753
/note="L2 repeat: matches 2647. .2749 of consensus"
repeat_region 11816. .12380
/note="LTR19B repeat: matches 1. .580 of consensus"
repeat_region 12519. .12813
/note="L1MD2 repeat: matches 6032. .6331 of consensus"
repeat_region 13559. .13726
/note="3 copies 56 mer 73% conserved"
repeat_region 13585. .13716
/note="22 copies 6 mer ctctct 67% conserved"
repeat_region 13586. .13717
/note="66 copies 2 mer tc 68% conserved"
repeat_region 13588. .13715
/note="4 copies 32 mer 78% conserved"
repeat_region 13589. .13724
/note="34 copies 4 mer cttt 77% conserved"
misc_feature complement(13622. .14142)
/note="match: GSS: Em:AQ592603"
repeat_region 14313. .14437
/note="LTR16C repeat: matches 257. .386 of consensus"
repeat_region 15015. .15110
/note="3 copies 32 mer 79% conserved"
repeat_region 15021. .15110
/note="45 copies 2 mer ac 80% conserved"
repeat_region 15026. .15109
/note="7 copies 12 mer 82% conserved"
repeat_region 15028. .15111
/note="14 copies 6 mer cacaca 82% conserved"
repeat_region 15029. .15108
/note="20 copies 4 mer acac 83% conserved"
repeat_region 15273. .15399
/note="L2 repeat: matches 2159. .2285 of consensus"
repeat_region 15420. .15721
/note="AluSx repeat: matches 1. .303 of consensus"
repeat_region 16333. .16396
/note="2 copies 32 mer 98% conserved"
repeat_region 16575. .16690
/note="L2 repeat: matches 2596. .2711 of consensus"
repeat_region 18049. .18169
/note="MIR repeat: matches 86. .211 of consensus"
repeat_region 18312. .18438
/note="MIR repeat: matches 138. .250 of consensus"
repeat_region 18585. .18776
/note="MER20 repeat: matches 7. .217 of consensus"
repeat_region 19211. .19290
/note="MIR repeat: matches 72. .154 of consensus"
repeat_region 20413. .20463
/note="MIR repeat: matches 90. .143 of consensus"
misc_feature complement(21005. .21672)
/note="match: GSS: Em:AQ748384"
misc_feature complement(21187. .21672)
/note="match: GSS: Em:AQ776209"
misc_feature complement(21278. .21663)
/note="match: GSS: Em:AQ136459"
repeat_region 21667. .21775
/note="MIR repeat: matches 73. .192 of consensus"
repeat_region 23577. .23646
/note="2 copies 35 mer 100% conserved"
repeat_region 24257. .24591
/note="L2 repeat: matches 2336. .2710 of consensus"
repeat_region 25580. .25611
/note="16 copies 2 mer tc 87% conserved"
repeat_region 26333. .26643
/note="AluY repeat: matches 1. .310 of consensus"
repeat_region 27603. .27684
/note="L2 repeat: matches 2661. .2739 of consensus"
repeat_region 27823. .28041
/note="MIR repeat: matches 38. .242 of consensus"
repeat_region 28129. .28259
/note="FLAM_C repeat: matches 1. .127 of consensus"
repeat_region 31203. .31356
/note="L1MB5 repeat: matches 6015. .6176 of consensus"
repeat_region 32224. .32522
/note="AluSx repeat: matches 1. .298 of consensus"

Query Match 70.5%; Score 748.2; DB 9; Length 130435;
Best Local Similarity 98.9%; Pred. No. 5e-131;
Matches 753; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
Qy 301 CGGAATCCAGAGCCGCGAGTCTACGCCCGCCCTCGGCCACCGACCGCGTGC 360
Db 128454 CTGCTTCCTCCAGCCGCGAGTCTACGCCCGCCCTCGGCCACCGACCGCGTGC 128395
Qy 361 GCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTAC 420
Db 128394 GCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTAC 128335
Qy 421 CGAGATCGACCTGCCGCCACCATCTCGTGTGACAGCGGGAGGAGCCCCACCTACCA 480
Db 128334 CGAGATCGACCTGCCGCCACCATCTCGTGTGACAGCGGGAGGAGCCCCACCTACCA 128275
Qy 481 GGGCCCTGCACCCCTCCAGCTTCGGGAGCCCCCGAGCAGCAGTGGAACTGAACCGGGAGTC 540
Db 128274 GGGCCCTGCACCCCTCCAGCTTCGGGAGCCCCCGAGCAGCAGTGGAACTGAACCGGGAGTC 128215
```



```
Best Local Similarity 98.1%; Pred. No. 2.7e-111;
Matches 683; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

Qy 359 CCGCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAG 418
    |||||
Db 693 CCGCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAG 634

Qy 419 CACGAGATCGACCTGCCGCCACCATCTCGTGTGCAGACGGGAGGAGCCCCACCTAC 478
    |||||
Db 633 CACGAGATCGA-CTGCCGCCACCATCTCGTGTGCAGACGG--AGGAGCCCCACCTAC 577

Qy 479 CAGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAG 538
    |||||
Db 576 CAGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAG 517

Qy 539 TCGGTGCGCGCACCCCCAAACAGAACCATCTTCGACAGTACCTGATGATGATGCCAGG 598
    |||||
Db 516 TCGGTGCGCGCACCCCCAAACAGAACCATCTTCGACAGTACCTGATGATGATGCCAGG 457

Qy 599 CTGGGCGGCCCTGCCCCCAGAGTAATCTGGGCATCAGGCCACCTGCTACGGCAGC 658
    |||||
Db 456 CTGGGCGGCCCTGCCCCCAGAGTAATCTGGGCATCAGGCCACCTGCTACGGCAGC 397

Qy 659 GCGGCGCATGGAGGGCGCGCCGCCACCTACAGCAGGATCATCGCCACTACCCGGGG 718
    |||||
Db 396 GCGGCGCATGGAGGGCGCGCCGCCACCTACAGCAGGATCATCGCCACTACCCGGGG 337

Qy 719 TCCTCTTCCAGCACCGAGCAGTGGCGCCCTCTTCTGCTGA-GGGGACCCGCT 777
    |||||
Db 336 TCCTCTTCCAGCACCGAGCAGTGGCGCCCTCTTCTGCTGAGGGGACCGGCT 277

Qy 778 CCACACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAA 837
    |||||
Db 276 CCCCCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAA 217

Qy 838 ACAGAAAGGACACCTCTCTAGGTCCCCAGGGGGCGGGCTGGGCTGCGTAGTGAA 897
    |||||
Db 216 ACAGAAAGGACACCTCTCTAGGTCCCCAGGGGGCGGGCTGCGTAGTGAA 157

Qy 898 AAGCGAAGCACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGCGGGGGCGCAGCAA 957
    |||||
Db 156 AAGCGAAGCACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAGCGGGGGCGCAGCAA 97

Qy 958 CGCATCGTGTGGCCCTCCCTCCCTCCCTGTGTATATAATTTACATGTGATGCTG 1017
    |||||
Db 96 CGCATCGTGTGGCCCTCCCTCCCTCCCTGTGTATATAATTTACATGTGATGCTG 37

Qy 1018 GTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1053
    |||||
Db 36 GTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1
```

```
RESULT 11
BC036995
LOCUS
DEFINITION Mus musculus, Similar to transmembrane, prostate androgen induced
RNA, clone IMAGE:5038092, mRNA.
ACCESSION BC036995
VERSION BC036995.1 GI:23331176
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1379)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
```

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, J., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 81 Row: i Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES
source

```
1..1379
/organism="Mus musculus"
/mol_type="mrna"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5038092"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MTV."
/clone_lib="NCI CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 352 a 390 c 349 g 288 t
ORIGIN
```

```
Query Match 55.0%; Score 584; DB 10; Length 1379;
Best Local Similarity 78.5%; Pred. No. 5.5e-100;
Matches 776; Conservative 0; Mismatches 175; Indels 37; Gaps 5;

Qy 82 TCTCCTCGAAACAGGCAATGGCGGAGCTGGAGTTGTTTCAGATCATCATCTGTTGT 141
    |||||
Db 120 TTTGTTCCCCAGCATGGAGATCAGGAGCTGGAGTTGTCGCAATCGTGGTCACTGTTGT 179

Qy 142 GGTGATGATGGTGTGTTGGTGTGATCACGTGCTGTGAGCCACTACAAGCTGTCTGC 201
    |||||
Db 180 AGTGATGATGGTGTGTTGGTGTGATTTACCTGCCTGCTGAGCCACTACAAGCTGTCTGC 239

Qy 202 ACGTCTCTTCATCAGCCGGCACAGCCAGGGCGGAGAGAGAAAGATGCCCTGTCTCAGA 261
    |||||
Db 240 CCGTCTCTTCATCAGCCGACACAGCCAGGCCAGGAGGAGAGACGATGGACTGTCTCGGA 299

Qy 262 AGGATGCTGTGGCCCTCGAGAGACACAGTGTGAGGCAACGGAATCCCAGAGCCGAGGT 321
    |||||
Db 300 AGGATGCTGTGGCCCTCGAGAGTACGGTGTGAGG---TGAATGCCGAGCCACAGGT 356

Qy 322 CTACGCCCCCGCTCGGCCACCGACCGCTGGCCCTGCGCCCTTCGCCCGAGCGGAGCG 381
    |||||
Db 357 CTATGCCCCCGCTCGGCCCACTGACCGACTCGCTGTGCCCCCTTCATCCAGCGG---- 411

Qy 382 CTTCACCGCTTCCAGCCACCTATCCGTACCTGTCAGCAGCAGATCGACCTGCCGCCAC 441
    |||||
Db 412 ----AGCCGATTCCAAACCCACCTACCCCTACCTGTCAGCAGCAAAATTGCCCTGCCACCC 467

Qy 442 CATCTCGTGTGAGAGCGGAGGAGCCCCACCTACAGGGCCCCCTGCACCTCCAGCT 501
    |||||
Db 468 CATCTCACTGTCTGATGGGAGGAGCCCCACCTACAGGGCCCCCTGCACCTCCAGCT 527

Qy 502 TCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCACCCCAACAG 561
```

||||| 528 ACGGACCTGAGCAACAGCTGGAGCTGAACGGGAATCTGTGGCGCACCCCTTAACCG 587
Qy 562 AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTGCCCCCCCAG 621
Db 588 GACCATCTTCGACAGTGACCTTATAGACAGCACCATGCTGGGGGGCCCTGTCCCCCAG 647
Qy 622 CAGTAACTCGGGCATCAGCGCCACCTGCTACAGCAGCGGTGGGCGCATGGAGGGCGCGC 681
Db 648 CAGTAACTCGGGCATCAGCGCCACCTGCTACAGCAGCGGTGGGCGCATGGAGGGCGCGC 707
Qy 682 GCCACCTACAGCGAGGTATCGGCCACTACCGGGGTCTCTCTCCAGCACCGACAGAG 741
Db 708 CCCCACCTACAGCGAGGTATCGGCCACTACCTGGCTCCTCTCCAGCACCGACGAAG 767
Qy 742 CAGTGGCGCCCTCCTTGTGTGAGGGGACCCGGCTCCACACACACATCGCGCCCT 801
Db 768 TAACGGGCCATCCTCCCTGCTAGAGGGGACCCGGCTCCATCACTCGCACATTGCCCCACT 827
Qy 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAAAGGATAAACAGAAAGGACACCCCTCTAGGG 861
Db 828 GGA-----GAACAAGAGAAAGGAGAAACAGAAAGGTACCCCTCTAGGA 872
Qy 862 TCCCAGGGGGCGGGCTGGG-----GCTCGTGTAGTGAAAGGCAGAACACTCGG 913
Db 873 GTGGGGCGGGCGCCTGTAGGCAAAACCGCAAAAAAAGGAAAAAACAACACTCGG 932
Qy 914 CGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCT 973
Db 933 CACTTCTTAAGAGAGAAAGAGAGAGGAAGTCAGGGGACACA--CAGGCCGAGTGGCCGTGT 990
Qy 974 CCCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAAGC 1033
Db 991 GGTAAGTCTCCTGTCTGTGTATAAATATTACATGTTCTGTGTGTCTGAATGCAGAAGC 1050
Qy 1034 TAAGAGAGCTTGCAAAAAAAGGAAAAA 1061
Db 1051 TCAAAAGCTTGCAAAAAAAGGAAAAA 1078

RESULT 12
AX392428
LOCUS AX392428 878 bp DNA linear PAT 23-MAR-2002
DEFINITION Sequence 12 from Patent WO0216416.
ACCESSION AX392428
VERSION AX392428.1 GI:19700744
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Lee,R.T., Landschulz,K.T., Kennedy,S.P., Thompson,J.F. and Turi,T.G.
TITLE Diagnosis and treatment of cardiovascular conditions
JOURNAL Patent: WO 0216416-A 12 28-FEB-2002;
THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US) ; PFIZER INC. (US)
FEATURES
source Location/Qualifiers
1..878
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
20..844
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD29011.1"
/db_xref="GI:19700745"
/translation="MGVNGTAAAGQPNVSCACNQRSLPFSPMEITELEFQIVIV
VVMVMVMITCLLSHYKLSARSFISRHSQARRRRDGLSSEGCLWPSESTVSGMPEP
QYAPPRPTDLAVPPFIQSRFPQTPYPLQHEIALPPTISLSDGEEPPPYQGPCTLQ
LRDPEQOLELNRESVRAPNRTIFDSDLIDSTMLGFCPPSSNSGISATCYSSGRME
GPPPTYSEVIGHYPGSSFOQSGNSPSSLEGLRLHSHIAPLENKEKEKQKHPL"
BASE COUNT 179 a 305 c 247 g 147 t

ORIGIN
Query Match 49.4%; Score 523.8; DB 6; Length 878;
Best Local Similarity 81.9%; Pred. No. 1.2e-88;
Matches 651; Conservative 0; Mismatches 117; Indels 27; Gaps 3;
Qy 82 TCTCTGCGAAACCCAGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 141
Db 94 TTTGTTCCCCAGCATGAGATCAGGAGCTGGAGTTCGTGCAATCGTGGTCACTCGTGT 153
Qy 142 GGTGATGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 201
Db 154 AGTGATGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 213
Qy 202 ACGGTCTTTCATCAGCGGCACAGCCAGGGCGGAGAGAGAGAGATGCCCTGTCTCAGA 261
Db 214 CCGTCTCTTCATCAGCGGCACAGCCAGGGCCAGGAGGAGAGACGATGGACTGTCTCGGA 273
Qy 262 AGGATGCCTGTGGCCTCGGAGAGCACAGTGTGAGGCAACGGAATCCCAGAGCCGCGAGST 321
Db 274 AGGATGCCTGTGGCCTCAGAGAGTACGGTGTGAGG--TGAATGCCGAGCCACAGST 330
Qy 322 CTACGCCCCCGCCTCGGCCACCGACCGCCTGGCCGCGCCCTTCGCCCGAGCGGGAGCG 381
Db 331 CTATGCCCCGCTCGGCCACTGACCGACTCGTGTGCCCCCTTCATCCAGCGG----- 385
Qy 382 CTTCCACCGCTTCCAGCCCCACTATCCGTACCTGACGACGAGATCGACCTGCCGCCAC 441
Db 386 ----AGCCGATTCCAAACCCACTACCCCTACCTGACGACGAAATTGCCCTGCCACCCAC 441
Qy 442 CATCTCGTGTGACAGCGGGAGGAGCCCCACCCCTACAGGGGCCCTGCACCTCCAGCT 501
Db 442 CATCTCACTGTGATGGGAGGAGCCCCCACCCCTACAGGGGCCCTGCACCTCCAGCT 501
Qy 502 TCGGGACCCCGAGCAGCAGCTGGAACCGGGAGTCCGTGCGCGCACCCCAACAG 561
Db 502 ACGGACCTTGAGCAACAGCTGGAGCTGAACCGGGAATCTGTGCGGCACCCCTAACCG 561
Qy 562 AACCATCTTCGACAGTGACCTGATGGATAGTGCAGGCTGGGCGGCCCTGCCGCCCAG 621
Db 562 GACCATCTTCGACAGTGACCTTATAGACAGCACCATGCTGGGGGGCCCTGTCCCCCAG 621
Qy 622 CAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGCGCGC 681
Db 622 CAGTAACTCGGGCATCAGCGCCACCTGTCTACAGCAGCGTGGGCGCATGGAGGGCGCGC 681
Qy 682 GCCACCTACAGCGAGGTGATCGGCCACTACCCGGGTCTCTCCAGCACACATCGCGCCCT 741
Db 682 CCCACCTACAGCGAGGTGATGGCCACTACCCCTGGTCTCTCTCCAGCACGACGAAG 741
Qy 742 CAGTGGCGCCCTCCTTGTGTGAGGGGACCCGGCTCCACCCACACACATCGCGCCCT 801
Db 742 TAACGGGCCATCCTCCTGTAGAGGGGACCCGGCTCCATCACTGACATTTGCCCCACT 801
Qy 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAAGAGAAAGGACACCCCTCTTAGGG 861
Db 802 GGA-----GAACAAGGAGAGGAGAAACAGAAAGGTCAACCCCTCTAGGA 846
Qy 862 TCCCCAGGGGGCGG 876
Db 847 GTGGGGCGCGGGCG 861

RESULT 13
AL837509/c
LOCUS AL837509 156698 bp DNA linear ROD 13-NOV-2002
DEFINITION Mouse DNA sequence from clone RP23-44L6 on chromosome 2, complete sequence.
ACCESSION AL837509
VERSION AL837509.10 GI:25045360
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 156698)

Almeida, J.
Direct Submission
Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Nov 15, 2002 this sequence version replaced gi:24414713.
Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-44L6 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6.

FEATURES
source

Location/Qualifiers
1..156698
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-44L6"
/clone_lib="RPCI-23"
39443 a 38596 c 39859 g 38800 t

BASE COUNT 39443 a 38596 c 39859 g 38800 t
ORIGIN

Query Match 40.8%; Score 433.4; DB 10; Length 156698;
Best Local Similarity 77.3%; Pred. No. 9e-72;
Matches 593; Conservative 0; Mismatches 136; Indels 38; Gaps 4;

Qy 307 CCCAGAGCGGAGGTCTACGCCCCCGCTCGGCCACCGACCGCTGGCGTGGCGCCCTT 366
Db |||||
CCTCCAGGCACAGGTCTATGCCCCCGCTCGGCCACCTGACCGACTCGTGTGCCCCCTT 135274

Qy 367 CGCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCACGAGAT 426
Db |||||
CATCCAGCGGAG-----CCGATTCACACCCACCTACCCCTACCTGCAGCACGAAAT 135223

Qy 427 CGACCTGCGGCCACCATCTCGCTGTTCAGACGGGAGGAGCCCCACCCCTACCAGGGCCC 486
Db |||||
TGCCCTGCCACCCACCATCTCACTGTGTGATGGGAGGAGCCCCACCCCTACCAGGGCCC 135163

Qy 487 CTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTGGC 546
Db |||||
CTGCACCTCCAGCTACGGGACCCCTGAGCAACAGCTGGAGCTGAACCGGGAATCTGTGCG 135103

Qy 547 CGCACCCCCAAACAGAACCATCTTCGACAGTGCACCTGATGGATAGTGCCAGGCTGGCGG 606
Db |||||
CGCACCCCCTAACCGGACCATCTTCGACAGTGCACCTTATAGACAGCACCATGCTGGGGG 135043

Qy 607 CCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACGTGTACGGCAGCGCGGCG 666
Db |||||
CCCCTGTCCCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGTACAGCAGCGGTGGCG 134983

Qy 667 CATGGAGGGCGCGCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGTCTCTCTT 726
Db |||||
CATGGAGGGCGCGCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGTGGTCTCTCTT 134923

Qy 727 CCAGCACCCAGCAGCAGTGGGGCGCCCTCTCTTGTGGAGGGGACCCGGCTCCACACAC 786
Db |||||
CCAGCACCCAGCAAAAGTAACGGGGCCATCTCTCCCTGTAGAGGGGACCCGGCTCCATCATC 134863

Qy 787 ACACATCGCGCCCCCTAGAGAGCGCAGCCATCTGTGAGCAAGAGAGGATAAAACAGAAAG 846
Db GCACATTGCCCCACTGGA-----GAACAAGGAGAGGAGAGAAACAGAAAGG 134818

Qy 847 ACACCCCTCTTAGGCTCCCGAGGGGGCGGGCTGGGCTGCGTAGGTGAAAAGGCAG-- 904
Db |||||
TCACCCCTCTTAGGAGTGGGGCGGGCGCTGTAGSCAAAACCGCAAAAAAAGAA 134758

Qy 905 -----AACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAG 954
Db |||||
AGGAAAAAAACACTCCGACTTCTTAAGAGAGAGAGAGAGAGGAGTACGGGGACACA- 134699

Qy 955 CAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATTTACATGTGTGT 1014
Db |||||
-CAGGCTGAGTGGCGGTGTGTGTGTCTCTCTCTGTGTATATAATTTACATGTGTGT 134640

Qy 1015 CTGGTCTGAATGCACAAGTAAAGAGAGCTTGCAAAAAAAGGAGGAGGAGGAGGAGGAGG 1061
Db |||||
GTGGTCTGAATGCAGAAGCTCAAAAAGCTTGCAAAAAAAGGAGGAGGAGGAGGAGGAGG 134593

RESULT 14
AC110189

LOCUS
DEFINITION
AC110189 175754 bp DNA linear HTG 21-OCT-2002
Mus musculus clone RP23-41204, WORKING DRAFT SEQUENCE, 15 ordered pieces.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC110189
AC110189.4 GI:24182070
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 175754)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-41204
Unpublished
2 (bases 1 to 175754)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hags, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,


```
QY 427 CGACCTGCGCCGCCACCATCTCGTGTCTAGACGGGGAGGAGCCGCCACCCCTACCGGGCC 486
Db 147035 TGCCCTGACACCCACCATCTCACTGTCTGATGGGAGGAGCCGCCACCCCTACCGGGCC 147094
QY 487 CTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTCGGTCCG 546
Db 147095 CTGCACCTCCAGCTACGGGACCCCTGAGCAACAGCTGGAGCTGAACCGGAACTCTGTCCG 147154
QY 547 CGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGTCGGGGCG 606
Db 147155 CGCACCCCTAACCGGACCATCTTCGACAGTGACCTTATAGACAGCACCATGTGTTGGGG 147214
QY 607 CCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACAGCAGCGGGCGG 666
Db 147215 CCCCTGTCCCCCAGCAGTAACCTCGGGCATCAGCGCCACCTGCTACAGCAGCGGGCGG 147274
QY 667 CATGGAGGGCGCGCGCCACCTACAGCGAGGTATCGGCCACTACCGGGGTCTCCTT 726
Db 147275 CATGGAGGGCGCGCGCCACCTACAGCGAGGTATCGGCCACTACCGGGGTCTCCTT 147334
QY 727 CCAGCACCGAGCAGCAGTGGGCGCCCTCTTCTGCTGGAGGGGACCCGGCTCCACCCAC 786
Db 147335 CCAGCACCGAGCAAGTAACGGGCCATCTCTCTGCTAGAGGGGACCCGGCTCCATC 147394
QY 787 ACACATCGCGCCCTAGAGAGCGGACGCCATCTGGAGCAAGAGAAAGGATAACAGAAAG 846
Db 147395 GCACATTGCCCCACTGGA-----GAACAAGGAGAGGAGAAACAGAAAG 147439
QY 847 ACACCTCTCTAGGTGCCCCAGGGGGCGGGCTGGGGCTGCGTAGGTGAAAGGCGAG-- 904
Db 147440 TCACCCCTCTAGGAGTGGGGCGCGGGCGCTGTAGGCAAAACCGCAAAAAAAGAA 147499
QY 905 -----AACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGGGGGCGGCGAG 954
Db 147500 AGGAAAAAACAACCTCCGCACTTCTTAAGAGAGAGAGAGAGGAAGTCAGGGGACACA- 147558
QY 955 CAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGCTGTGTATAAATATTATCATGTGT 1014
Db 147559 -CAGGCTAGTGGCCGTGTGGTAGTCTCTCTGCTGTGTATAAATATTATCATGTGT 147617
QY 1015 CTGCTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAGGAGGAGGAGGAGGAGG 1061
Db 147618 GTGGTCTGAATGCAGAAGCTCAAAAGCTTGCAAAAAAGGAGGAGGAGGAGGAGGAG 147664
```

```
RESULT 15
AL837520/c
LOCUS AL837520 176821 bp DNA linear HTG 01-JUN-2003
DEFINITION Mus musculus chromosome 2 clone RP23-41204, *** SEQUENCING IN
PROGRESS ***
ACCESSION AL837520
VERSION AL837520.25 GI:31335607
KEYWORDS HTG; HTGS PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 176821)
AUTHORS Bates,K.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
COMMENT On Jun 2, 2003 this sequence version replaced gi:30962480.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
```

```
----- Project Information
Center project name: bm41204
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 99% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 176788 bases at least Q40
Consensus quality: 176819 bases at least Q30
Consensus quality: 176821 bases at least Q20
Insert size: 176821; sum-of-contigs
Insert size: 180785; 10.3% error; agarose-fp
Quality coverage: 9.74x in Q20 bases; sum-of-contigs Quality
coverage: 9.53x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 176821: contig of 176821 bp in length.
-----
FEATURES
source
1.176821
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-41204"
/clone_lib="RPCI-23"
1.176821
/note="assembly_fragment:00029
clone_end:SP6
vector_side:right"
misc_feature
1.176821
/note="assembly_fragment:00029
clone_end:SP6
vector_side:right"
BASE COUNT 43349 a 44412 c 45766 g 43294 t
ORIGIN
Query Match 40.8%; Score 433.4; DB 2; Length 176821;
Best Local Similarity 77.3%; Pred. No. 8.9e-72;
Matches 593; Conservative 0; Mismatches 136; Indels 38; Gaps 4;
QY 307 CCCAGAGCCGAGGTCCTAGCGCCCGCCCTCGGCCACGACCGCCCTGGCCGTCGCCCTT 366
Db 28831 CCTCCAGCCACAGGTCATGCGCCCGCCCTCGGCCACTGACCGACTCGTGTGCCCCCTT 28772
QY 367 CGCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCCGCCCTATCCGTACCTGCAGCAGGAT 426
Db 28771 CATCCAGCGGAG-----CCGATTCCAACCCCTACCCCTACCTGCAGCAGCAAT 28721
QY 427 CGACCTGCCGCGCCACCATCTCGCTGTCTGATGGGGAGGAGCCCGCCACCCCTACCGGGCC 486
Db 28720 TGCCCTGCCACCCACCATCTCACTGTCTGATGGGGAGGAGCCCGCCACCCCTACCGGGCC 28661
QY 487 CTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTCCG 546
Db 28660 CTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTCCG 28601
QY 547 CGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGGCGG 606
Db 28600 CGCACCCCTAACCGGACCATCTTCGACAGTGACCTTATAGACAGCACCATGTGTTGGGG 28541
QY 607 CCCCTGCCCCCAGCAGTAACCTCGGGCATCAGCGCCAGCTGTACGGCAGCGGGCGG 666
Db 28540 CCCCTGTCCCCCAGCAGTAACCTCGGGCATCAGCGCCAGCTGTACAGCAGCGGTGGCG 28481
QY 667 CATGGAGGGCGCGCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGTCTCTCTT 726
Db 28480 CATGGAGGGCGCGCGCCACCTACAGCGAGGTCTATCGGCCACTACCCGGGTCTCTCTT 28421
QY 727 CCAGCACCGAGCAGTGGGGCGCCCTCTTGTGAGGGGACCGGGCTCCACACAC 786
```


Db 28420 CCAGCACAGCAAAGTAACGGGCCATCTCCCTGTAGAGGGACCCGGCTCCATCACTC 28361

Qy 787 ACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGG 846

Db 28360 GCACATTGCCCCCACTGGA-----GAACAAGGAGAGGAGAAACAGAAAGG 28316

Qy 847 ACACCCCTCTTAGGGTCCCCAGGGGGCCGGCTGGGGCTGGGTAGGTGAAAAGGCAG-- 904

Db 28315 TCACCCCTCTAGGAGTGGGGGCCGGGGCCCTGTAGGCAAAACCGCAAAAAAGAA 28256

Qy 905 -----AACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGAGGGGGGGGGCGCAG 954

Db 28255 AGGAAAAAAACACTCCGCACCTCTTAAGAGAGAGAGAGAGGAAGTCAGGGGACACA- 28197

Qy 955 CAACGCATCGTGTGGCCCTCCCCTCCACCTCCCTGTGTATAAATATTTACATGTATGT 1014

Db 28196 -CAGGCTGAGTGGCCGTGTGGTAGTTCTCCTGTCTGTGTATAAATATTTACATGTCTGT 28138

Qy 1015 CTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAGAAAAA 1061

Db 28137 GTGGTCTGAATGCAGAAGCTCAAAAAGCTTGCAAAAAAGAAAAAGA 28091

Search completed: December 8, 2003, 03:10:56
Job time : 4140 secs

THIS PAGE BLANK (USPTO)

Badner, Judith A.
Goldin, Lynn R.
Berrettini, Wade H.
Yoshikawa, Takeo
Sanders, Alan R.
Esterling, Lisa E.
TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,952A
FILING DATE: 19-Apr-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/029,278
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: PCT/US97/19381
FILING DATE: 28-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-297100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1...921
OTHER INFORMATION: Clone 22 coding region
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-091-952A-7

Query Match 32.9%; Score 349; DB 4; Length 921;
Best Local Similarity 73.7%; Pred. No. 9.7e-73;
Matches 474; Conservative 0; Mismatches 160; Indels 9; Gaps 2;
Qy 105 CGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGGTGGTGG 164
Dy 182 CGGAGCTGGAGTTTCGCCCAATCATCATCATCGTGGTGGTGGTGGTGGTGG 241
Qy 165 TGATCAGCTGCCTGCTGAGCCACTACAAGTGTCTGCAAGGATGCCTGTGGCCCTCGGAC 224
Dy 242 TCATCGTCTGCCTGCTGAACCACTACAAGTCTCCACGGGTCTTCATCAACCGCCGA 301
Qy 225 GCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGACA 284
Dy 302 ACCAGAGCCGGAGGCGGAGGACGGCTGCCGAGGAGGGTGCCTGTGGCCCTTCAGACA 361
Qy 285 GCACAGTGTGAGCAACGGAATCCAGAGCCGAGGTCTACGCCCCCGCTCGGCCACCG 344
Dy 362 GCGCGCACCGCGGTGGG-----CGCCTCGGAGATCATGATGCCCCCGGTCCAGGG 415
Qy 345 ACCGCTGGCCGTGTCGCCCCCTTCGCCAGGGAGGGCTTCACCGCTTCCAGCCCACT 404

416 ACAGGTTACAGCGCGCTCTTCCATCCAGAGGGATCGCTTCAGCCGCTTCCAGCCCACT 475
Qy 405 ATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGTGTGACAGCGGGAGG 464
Dy 476 ACCCTATGTGCAGCAGCAGATGATCTTCTCTCCACCATCTCCCTGTCCGACGGTGAAG 535
Qy 465 AGCCCCCACCCTACAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524
Dy 536 AGCCACCTCCTTACCAGGGCCCTGCACCCCTGCAGCTCCGGGACCCCTGAACAGCAGATGG 595
Qy 525 AACTGAACCGGGAGTCGTGCGCGCACCCCCCAACAGAACCATCTTCGACAGTGACCTGA 584
Dy 596 AACTCAACCGAGAGTCGTGAGGGCCCCCAACCCAGAACCATATTTGACAGTGATTAA 655
Qy 585 TGGATAGTCCCAGG---CTGGGGGGCCCCCTGCCCCCCCAGCAGTAACCTCGGGCATCAGCG 641
Dy 656 TAGACATTGCTATGTATAGCGGGGGTCCATGCCCAACCCAGCAGCAACTCGGGCATCAGTG 715
Qy 642 CCACGTGCTACGGCAGCGCGGGCGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTCA 701
Dy 716 CAAGCACCTGCAGCAGTAACGGGAGGATGGAGGGGCCACCCCCACATACAGCGAGGTGA 775
Qy 702 TCGGCCACTACCGGGGTCTCTCTTCCAGCACCCAGCAGCAGCAGCAG 744
Dy 776 TGGGCCACCACCCAGGCGCTCTTTCTCTCCATCACCAGCGCGCAG 818

RESULT 4

US-09-091-952A-6
; Sequence 6, Application US/09091952A
; Patent No. 6458532
; GENERAL INFORMATION:
; APPLICANT: Detera-Wadleigh, Sevilla D.
; Gershon, Elliot S.
; Badner, Judith A.
; Goldin, Lynn R.
; Berrettini, Wade H.
; Yoshikawa, Takeo
; Sanders, Alan R.
; Esterling, Lisa E.
; TITLE OF INVENTION: Chromosomal Markers and Diagnostic Tests for Manic-Depressive Illness
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,952A
; FILING DATE: 19-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,278
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: PCT/US97/19381
; FILING DATE: 28-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 015280-297100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; TELEX: <Unknown>


```

; EARLIER FILING DATE: 1993-10-07
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 76
; LENGTH: 53500
; TYPE: DNA
; ORGANISM: Streptomyces lavendulae
US-09-266-965-76

```

	Query Match	4.5%;	Score 47.4;	DB 4;	Length 53500;
	Best Local Similarity	47.2%;	Pred. No. 0.12;		
	Matches 176;	Conservative	0;	Mismatches 196;	Indels 1; Gaps 1;
QY	441	CCATCTCGTGT	CAGACGGGAGGAGCC	CCCCACCCCTAC	CAGGGCCCTGCACCCCTCCAGC 500
Db	13517	CCACGT	CGGTGACGAGGAAGCG	CGGGCGGTGCTCT	CACCGTCTCCCGGGGC 13576
QY	501	TTCGGGACCC	CGAGCAGAGCTGGA	ACTGAACCGGAGT	CGGTGCGCGCACCCCAACA 560
Db	13577	CTGGCCGCG	CGGTGCCGAGGAGCG	CGGCTCCAGTCGG	CCGTGGCCTCGCCCCGGTCA 13636
QY	561	GAACCATCT	TGCACAGTGACCT	GATGATGATGCC	AGGTGTCAGGCTGGCGGGCCCCCTGCCCCCCCA 620
Db	13637	TCCAGT	CCGCGACATCGT	GCTCGCCGCCCGGAC	CACCGCGGGCCACCCAGCAGGTGCA 13696
QY	621	GCAGTAACT	CGGGCATCAGCGCC	ACGTGCTACGGC	AGCGGGGCGCATGGAGGGCGCG 680
Db	13697	GGCCACGT	CCAGCAGGTGCC	AGCCAGGTCGAG	CAGCGCGCGCCCCCGGCAGCCGCG 13756
QY	681	CGCCCACT	ACAGCGAGGTCA	TCCGCCACTAC	CCGGGTCTCTTCCAGCACCCAGCAGA 740
Db	13757	GGTCGAC	AAACCAACCCGGT	GCGTGC	GGGATGCCGTGGCCCGGATCCAGC-TCAGCCCG 13815
QY	741	GCAGTGG	CGCCCTCTCTT	GCTGGAGGGAC	CCGGCTCCACACACACATCGCGCCCC 800
Db	13816	ACACTG	CGCAGCGT	GCCCCAGCGAGG	CAGCAGTCCGCCCAGCGGCGCAGACATCGGTGCGG 13875
QY	801	TAGAGAG	CGCAGC	813	
Db	13876	TGCCGG	CGGCGC	13888	

RESULT 11
US-09-252-991A-4586
; Sequence 4586, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

	Query Match	4.4%	Score 46.2;	DB 4;	Length 561;
	Best Local Similarity	51.2%;	Pred. No. 0.056;		
	Matches 108;	Conservative	0;	Mismatches 103;	Indels 0; Gaps 0
Qy	412	CCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGCTGTCTAGACGGGGAGGACCC	471		
Db	175	CCTGCTGCACGATATCGGCCACCTCTACGAAGACCCGGCGCAGATCGACGAGGAGGACCT	234		
Qy	472	ACCCTACCAAGGCCCTTGCAACCTTCAGCTTCGGGACCCCGAGCAGCAGCTGGAAC	531		

Db	235	GCGCCACGAGGAATTCGGCGCAGCCCTGTCGGCGAGCTGTTCCACAGTCTGGTCTGGCA	294
Qy	532	CCGGGAGTCGGTGC GCGCACCCCCAAACAGAACCATCTTCGACAGTCACCTGATGGATAG	591
Db	295	GCCGGTGC GCGCTGCACGTATCGGCCAAGCGCTTCTCTGCGCGGTGACCCGAGCTACCA	354
Qy	592	TGCCAGGCTGGCGGCGCCCTGCCCCCCCCAGC	622
Db	355	CGCCAGCCTGTGCGCGGCGCTCCCGGCACAGC	385

RESULT 12
US-09-252-991A-4391
; Sequence 4391, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4391
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4391

```
Query Match      4.4%; Score 46.2; DB 4; Length 1203;
Best Local Similarity 51.2%; Pred. No. 0.071;
Matches 108; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
```

Qy	412	CCTGCAGCACGAGATCGACCTGCGGCCACCATCTCGCTGTCAGACGGGGAGGAGCCCC	471
Db	783	CCTGCTGCACGATATCGGCCACCTCTACGAAGACCCGGCGCAGATCGACGAGGAGGACCT	842
Qy	472	ACCCATACAGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAAGTAA	531
Db	843	CGCCACAGGAATTGGCGCACGCCTGCTGCGCGAGCTGTTCCAGAGTCGGTCTGGCA	902
Qy	532	CCGGGAGTCGTCGCGCACCCCCAAACAGAACCATCTTCGACAGTGACCTGATGGATAG	591
Db	903	GCCGGTCGCCCTGCACGATCGGGCCAAGCGCTTCCTCTGCGCGTGGACCCGAGCTACCA	962
Qy	592	TGCCAGGCTGGCGGCCCTCTGCCCCCCCCAGC	622
Db	963	CGCCAGCCTGTGCGCGGCCCTCCCGGCACAGC	993

RESULT 13
US-09-252-991A-4200/C
; Sequence 4200, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

; FILE REFERENCE: 10/196.138
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4200
; LENGTH: 1611

```



```

; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4200

Query Match      4.4%; Score 46.2; DB 4; Length 1611;
Best Local Similarity 51.2%; Pred. No. 0.078;
Matches 108; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 412 CCTGCAGCAGATCGACCTGCGCGCCACCACCTCTCGCTGTACAGCGGGAGGCCCCC 471
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1474 CCTGCTGCAGATATCGGCCACCTTACGAAGACCCGGCGCAGATCGACGAGGAGCCT 1415

Qy 472 ACCCTACAGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGGAACCTGAA 531
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1414 GCGCCACGAGGAATTTCGGCGCACGCTGCTGCGGAGCTGTTCCACAGATCGGTCTGGCA 1355

Qy 532 CCGGAGTCGGTGCAGCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAG 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1354 GCCGGTGGCCTGCACGTATCGGCCAAGCGCTTCCTCTGCGCGTGGACCCGAGCTACCA 1295

Qy 592 TGCCAGGCTGGCGGCCCTGCCCCCCCCCAGC 622
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1294 CGCCAGCTGTGCGCGGCTCCCGGCACAGC 1264
```

```

RESULT 14
US-08-785-420-1/c
; Sequence 1, Application US/08785420
; Patent No. 6001976
; GENERAL INFORMATION:
; APPLICANT: MacLennan, David H
; APPLICANT: O'Brien, Peter J.
; TITLE OF INVENTION: DIAGNOSIS FOR PORCINE MALIGNANT
; TITLE OF INVENTION: HYPERTHERMIA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: P.O. Drawer 34009
; CITY: Charlotte,
; STATE: No. 6001976th Carolina 28234
; COUNTRY: U.S.A.
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/785,420
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,388
; FILING DATE:
; APPLICATION NUMBER: US 08/030,159
; FILING DATE: 15-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Layton, Jr., Samuel G.
; REGISTRATION NUMBER: 22807
; REFERENCE/DOCKET NUMBER: 3477-73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 704-377-1561
; TELEFAX: 704-334-2014
; TELEX: 57-5102
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15378 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```

; IMMEDIATE SOURCE:
; CLONE: Porcine RYR1 Gene
; POSITION IN GENOME:
; UNITS: bp
US-08-785-420-1

Query Match      4.4%; Score 46.2; DB 3; Length 15378;
Best Local Similarity 50.7%; Pred. No. 0.15;
Matches 111; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 320 GTCTACGCCCGCCCTCGGCCACCGACCGCTGGCGTGGCGCGCCCTTCGCCCCAGCGGGAG 379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13329 GTCCGCTCGCCCGCGGCCAGCCGGCTGCTCGCCGTGCACCTCGTCCCGCTGGGATC 13270

Qy 380 CGCTTCCACCGCTTCCAGGCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCC 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13269 GGGCATGCCCGCCAGGAGCTCGGTACCGTCACTTCTTGGCACCTCCACCGCCGCC 13210

Qy 440 ACCATCTCGCTGTCAGACGGGGAGGAGCCCCACCCCTTACCAGGGCCCCCTGCACCTCCAG 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13209 ACCAAGAGCGAGCCCCCAGAGCAGCGCGAGCGCGCTGCGCGCCCCCGCGCGCCGCGC 13150

Qy 500 CTTCGGGACCCCGAGCAGAGCTGGAACCTGAACCGGGAG 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13149 CCGCGGTGCGCCAGCGCCGCCAGAGCAGCGCGGCCAG 13111
```

```

RESULT 15
US-09-252-991A-692
; Sequence 692, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 692
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-692

Query Match      4.3%; Score 46; DB 4; Length 3756;
Best Local Similarity 46.5%; Pred. No. 0.11;
Matches 148; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 377 GAGCGCTTCCACCGCTTCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCG 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3232 GACACCTCCGCCAGCACCTGGCGCTGGCAGCCCCGGCGCGCAGAAGCTCGGCATCAAT 3291

Qy 437 CCCACCATCTCGTGTCAAGCGGGAGGAGCCCCACCTACAGGGCCCCCTGCACCTC 496
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3292 CCGGCGTGTGCACACCTTCCAGCGCGGCGAGCGATCCGCGACGCTTCTTCCGCTCC 3351

Qy 497 CAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGAGTGGTGGCGCACCCCCA 556
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3352 GCGGCATGCAGCCGCGGTGCGCTTCGAACCTCAAGCGGTGACCATGGACGCGCGATC 3411

Qy 557 AACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCAGGCTGGGCGGCCCTGCCCC 616
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3412 AGCCAGTTTCATCTCGACCTCGACGGCCAGCAGCTGACCTACGACCGGCCCGCCGCGC 3471

Qy 617 CCCAGCAGTAACTCGGGCATCAGCGCCACGCTGCTACGCGAGCGGGCGCATGGAGGG 676
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3472 CCGGTGGCATGCAGTGGCCGAGCGCCAAACGGCCTCGCGTGGTCCGCTGACGGTGACG 3531
```

Qy 677 CCGCGCCCGCCACCTACAGC 694
Db 3532 CCGCGCCCGCCACGAGCGGC 3549

Search completed: December 8, 2003, 03:57:31
Job time : 84 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 01:15:30 ; Search time 2637 Seconds
(without alignments)
9778.934 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tcctccttggttcgggtga.....cttcgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				EST:*			
1:	em_estba:	*		1:	em_estba:	*	
2:	em_esthum:	*		2:	em_esthum:	*	
3:	em_estin:	*		3:	em_estin:	*	
4:	em_estmu:	*		4:	em_estmu:	*	
5:	em_estov:	*		5:	em_estov:	*	
6:	em_estpl:	*		6:	em_estpl:	*	
7:	em_estro:	*		7:	em_estro:	*	
8:	em_htc:	*		8:	em_htc:	*	
9:	gb_est1:	*		9:	gb_est1:	*	
10:	gb_est2:	*		10:	gb_est2:	*	
11:	gb_htc:	*		11:	gb_htc:	*	
12:	gb_est3:	*		12:	gb_est3:	*	
13:	gb_est4:	*		13:	gb_est4:	*	
14:	gb_est5:	*		14:	gb_est5:	*	
15:	em_estfun:	*		15:	em_estfun:	*	
16:	em_estom:	*		16:	em_estom:	*	
17:	em_gss_hum:	*		17:	em_gss_hum:	*	
18:	em_gss_inv:	*		18:	em_gss_inv:	*	
19:	em_gss_pln:	*		19:	em_gss_pln:	*	
20:	em_gss_vrt:	*		20:	em_gss_vrt:	*	
21:	em_gss_fun:	*		21:	em_gss_fun:	*	
22:	em_gss_mam:	*		22:	em_gss_mam:	*	
23:	em_gss_mus:	*		23:	em_gss_mus:	*	
24:	em_gss_pro:	*		24:	em_gss_pro:	*	
25:	em_gss_rod:	*		25:	em_gss_rod:	*	
26:	em_gss_phg:	*		26:	em_gss_phg:	*	
27:	em_gss_vrl:	*		27:	em_gss_vrl:	*	
28:	gb_gss1:	*		28:	gb_gss1:	*	
29:	gb_gss2:	*		29:	gb_gss2:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	889.4	83.8	1201	9	AL517150
C 2	878.4	82.8	1009	9	AL578575
C 3	858.8	80.9	1007	9	AL558882
C 4	844.2	79.6	951	9	AL558881

C 5	794.4	74.9	874	13	BX362396
C 6	776	73.1	945	13	BU539219
C 7	773	72.9	967	13	BQ641849
C 8	768.4	72.4	782	12	BQ015170
C 9	724.4	68.3	729	13	BQ575741
C 10	720.2	67.9	730	12	BM677602
C 11	709.8	66.9	728	13	BU683523
C 12	697.8	65.8	850	13	BU602918
C 13	680.2	64.1	890	13	BQ690750
C 14	652	61.5	655	13	BQ691705
C 15	644.2	60.7	693	9	AI761441
C 16	636.2	60.0	646	13	BU859841
C 17	635.2	59.9	964	13	BU859860
C 18	612.4	57.7	629	13	BU730650
C 19	609.2	57.4	618	14	CD367193
C 20	609	57.4	609	13	BQ636742
C 21	605	57.0	626	12	BM974296
C 22	599	56.5	602	14	CA431191
C 23	594.8	56.1	1046	12	BM922276
C 24	592.8	55.9	633	12	BM714472
C 25	591	55.7	651	14	CB554226
C 26	587.2	55.3	1079	11	BC023092
C 27	585	55.1	588	13	BU624784
C 28	578	54.5	844	13	BQ686793
C 29	578	54.5	952	13	BU157959
C 30	573	54.0	973	13	BU169156
C 31	572.2	53.9	641	9	AW071693
C 32	569.4	53.7	938	13	BU157842
C 33	562.2	53.0	570	13	BQ575582
C 34	562.2	53.0	916	13	BQ954555
C 35	559.2	52.7	674	9	AI972096
C 36	559	52.7	559	10	BE855409
C 37	555.8	52.4	619	9	AI742327
C 38	548.4	51.7	551	12	BM141979
C 39	534	50.3	547	12	BM676516
C 40	534	50.3	552	12	BM713900
C 41	534	50.3	563	14	CB049800
C 42	528.6	49.8	588	9	AI377498
C 43	523.8	49.4	1207	11	AK008976
C 44	515.8	48.6	634	9	AI826012
C 45	502	47.3	518	9	AI885001

ALIGNMENTS

RESULT 1
AL517150/c
LOCUS
DEFINITION AL517150 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CSODA008YB23 3-PRIME, mRNA sequence.
ACCESSION AL517150
VERSION AL517150.2 GI:30492472
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 13, 2001 this sequence version replaced gi:12780643.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9945.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODA008CAL2NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :

AL517150 1201 bp mRNA linear EST 09-MAY-2003
AL517150 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CSODA008YB23 3-PRIME, mRNA sequence.

http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DA008CA12NP1.

FEATURES

Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA008VB23"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 177 a 372 c 401 g 214 t 37 others
ORIGIN
Query Match 83.8%; Score 889.4; DB 9; Length 1201;
Best Local Similarity 97.2%; Pred. No. 6.3e-157;
Matches 913; Conservative 1; Mismatches 24; Indels 1; Gaps 1;

QY 82 TCTCCTGCAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 141
Db 939 TTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 880
QY 142 GGTGATGATGGTGGTGGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGC 201
Db 879 GGTGATGATGGTGGTGGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGC 820
QY 202 ACGTCTTTCATCAGCCGGCACAGCCAGGGGGCGGAGAGAGATGCCCTGTCTCAGA 261
Db 819 ACGTCTTTCATCAGCCGGCACAGCCAGGGGGCGGAGAGAGATGCCCTGTCTCAGA 760
QY 262 AGGATGCTGTGGCCCTCGGAGAGACAGTGTGAGGCAACGGAAATCCAGAGCCGAGGT 321
Db 759 AGGATGCTGTGGCCCTCGGAGAGACAGTGTGAGGCAACGGAAATCCAGAGCCGCA-GT 701
QY 322 CTACGCCCCGCTCGGCCACCGACCCGCTGGCCGCTTGGCCGCTTGGCCGAGGGAGCG 381
Db 700 CTACGCCCCGCTCGGCCACCGACCCGCTGGCCGCTTGGCCGCTTGGCCGAGGGAGCG 641
QY 382 CTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCAC 441
Db 640 CTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCAC 581
QY 442 CATCTCGTGTGACAGCGGGAGGAGCCGCCACCCCTACCAGGGCCCTGCACCCCTCCAGCT 501
Db 580 CTTCTCGTGTGACAGCGGGAGGAGCCGCCACCCCTACCAGGGCCCTGCACCCCTCCAGCT 521
QY 502 TCGGGACCCGAGCAGCAGCTGGAACCGGGAGTGGTGGCGGCGACCCCAACAG 561
Db 520 TCGGGACCCGAGCAGCAGCTGGAACCGGGAGTGGTGGCGGCGACCCCAACAG 461
QY 562 AACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGCGGCGCCCTGCCCCGCCAG 621
Db 460 AACCATCTTCGACAGTACCTGATGATAGTGCAGGCTGGCGGCGCCCTGCCCCGCCAG 401
QY 622 CAGTAACCTCGGCGATCAGCGCCACGTGCTACGGCAGCGGGCGGCGATGGAGGGCGGCC 681
Db 400 CAGTTACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCGATGGAGGGCGGCC 341
QY 682 GCCCACCTACGCGAGGTATCGGCCACTACCGGGGGTCTCTCCATCCAGCACGACGAG 741
Db 340 GCCCACCTACGCGAGGTATCGGCCACTACCGGGGGTCTCTCCATCCAGCACGACGAG 281
QY 742 CAGTGGGCGCCCTCTTGTGGAGGGGACCGGGCTCCACACACACATCGCGGCCCT 801
Db 280 CAGTGGGCGCCCTCTTGTGGAGGGGACCGGGCTCCCCACACACAGCTCGCGGCCCT 221
QY 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGATAAACAGAAAGGACACCCCTCTCTAGGG 861
Db 220 AGAGAGCGCAGCCCTTCTGGAGCAAGAGAGATAAACAGAAAGGACACCCCTCTCTAGGG 161

QY 862 TCCCCAGGGGGCGGCTGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTT 921
Db 160 TCCCCAGGGGGCGGCTGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTT 101
QY 922 AGAAGAGGAGTGAGAGAAAGCGGGGGCGGCGAGCAACGATCGTGTGGCCCTCCCTCC 981
Db 100 AGAAGAGGAGTGAGAGAAAGCGGGGGCGGCGAGCAACGATCGTGTGGCCCTCCCTCC 41
QY 982 ACCTCCCTGTGTATAATAATTACATGTGATGTCTGGTC 1020
Db 40 ACCNCCCVRTGTATAATAATTAAACNTGTGATGTCTGGTC 2

RESULT 2

AL578575/c
LOCUS
DEFINITION
AL578575 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK001YC24 3-PRIME, mRNA sequence.
ACCESSION
AL578575
VERSION
AL578575.2 GI:31316780
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1009)
AUTHORS
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
COMMENT
On Feb 16, 2001 this sequence version replaced gi:12942781.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9945.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DK001BB12NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DK001BB12NP1.

FEATURES

Location/Qualifiers
1. .1009
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK001YC24"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 160 a 293 c 344 g 201 t 11 others
ORIGIN
Query Match 82.8%; Score 878.4; DB 9; Length 1009;
Best Local Similarity 96.3%; Pred. No. 7.5e-155;
Matches 896; Conservative 11; Mismatches 22; Indels 1; Gaps 1;

QY 82 TCTCTGCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 141
Db 929 TTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCGTGT 870
QY 142 GGTGATGATGGTGGTGGTGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGC 201
Db 869 GGTGATGATGGTGGTGGTGATCAGCTGCCTGCTGAGCCACWACAAGCTGTCTGC 810
QY 202 ACGTCTTTCATCAGCCGGCACAGCCAGGGGGCGGAGAGAGATGCCCTGTCTCAGA 261

Db 809 ACGGTCTCTTTCAGCCGGCACAGCCAGGGGGGAGGAGAGAGATGCCCTGTCTCAGA 750

Qy 262 AGGATGCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAACGGGAATCCAGAGCCGCGAGT 321

Db 749 AGGATGCTGTGGCCCTCGGAGAGCACAGTGTTCAGGCAACGGGAATCCAGAGCCGCA-GT 691

Qy 322 CTACGCCCCCTCGGCCACCCGACCGCCTGGCCGTGGCCCTTCGCCCCAGCGGGAGCG 381

Db 690 CTACGCCCCCTCGGCCACCCGACCGCCTGGCCGTGGCCCTTCGCCCCAGCGGGAGCG 631

Qy 382 CTTCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCAC 441

Db 630 CTTCACCGCTTCCAGCCCCACCTATCCGTACCTGCAGCAGAGATCGTCTGCCGCCAC 571

Qy 442 CATCTCGCTGTACAGACGGGGAGAGCCCCACCTACAGGGGCCCTGCACCCCTCCAGCT 501

Db 570 CTTCCTCGCTGTTCAGACGGGGAGAGCCCCACCTACAGAGGCCCTGCACCCCTCCAGCT 511

Qy 502 TCGGACCCCGAGCAGCAGCTGGAAGTGAACCGGGAGTTCGGTGGCGGCACCCCAACAG 561

Db 510 TCGAGACCCCGAGCAGCAGCTGGAAGTGAACCGGGAGTTCGGTGGCGGCACCCCAACAG 451

Qy 562 AACCATCTTCAGCAGTACCTGTATGATAGTGCAGGCTGGGGGCCCTGCCGCCCCAG 621

Db 450 AACCATCTTCAGCAGTACCTGTATGATAGTGCAGGCTGGGGGCCCTGCCGCCCCAG 391

Qy 622 CAGTAACCTCGGGCATCAGCGGCACCTGTACGGCAGCGGGCGCATGGAGGGCGCC 681

Db 390 CAGTAACCTCGAGCATCAGCGGCACCTGTACRGACRGCGGGCGCATGGAGRGCGCC 331

Qy 682 GCCACCTACAGCAGGTCATCGGCCACTACCGGGTCTCTTCAGCAGCAGCAGAG 741

Db 330 GCCACCTACAGCAGGTCATCGGCCACTACCGGGTCTCTTCAGCAGCAGCAGAG 271

Qy 742 CAGTGGGCGCCCTCTTGTGGAGGGGACCGGCTCCACACACACACATCGCGCCCT 801

Db 270 CAGTGGGCGCCCTCTTGTGGAGGGGACCGGCTCCACACACACACATCGCGCCCT 211

Qy 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGG 861

Db 210 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGG 151

Qy 862 TCCCGAGGGGGCGGCTGGGCTGGGTAGGTGAAAGGAGAGAACTCCGCGCTTCTT 921

Db 150 TCCCGAGGGGGCGGCTGGGCTGGGTAGGTGAAAGGAGAGAACTCCGCGCTTCTT 91

Qy 922 AGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCCC 981

Db 90 AGTAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCCC 31

Qy 982 ACCTCCCTGTGTATAAATATTACATGTGA 1011

Db 30 ACCTCCCTGTGTATAAATATTACATRTA 1

RESULT 3

AL558882 1007 bp mRNA linear EST 31-MAY-2003

LOCUS

DEFINITION AL558882 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

Homo sapiens cDNA clone CS0DJ015YF12 5-PRIME, mRNA sequence.

AL558882

VERSION AL558882.2 GI:31283015

KEYWORDS

SOURCE EST.

ORGANISM Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1007)

Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

On Feb 15, 2001 this sequence version replaced gi:12903838.

Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9945.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015DC06QP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015DC06QP1.

FEATURES
source

1..1007
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YF12"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 218 a 325 c 292 g 151 t 21 others
ORIGIN

Query Match 80.9%; Score 858.8; DB 9; Length 1007;
Best Local Similarity 98.8%; Pred. No. 3.5e-151;
Matches 891; Conservative 5; Mismatches 2; Indels 4; Gaps 3;

Qy 160 GGTGGTGATCACGTGCTCTGTGAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCG 219

Db 62 GGATGTGATCACGTGCTCTGTGAGCCACTACAAGCTGTCTGACGGTCTTTCATCAGCCG 121

Qy 220 GCACAGCCAGGGGGGAGGAGAGAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTC 279

Db 122 GCACAGCCAGGGGGGAGGAGAGAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTC 181

Qy 280 GGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCCGCGAGTCTACGCCCGCCCTCGGCC 339

Db 182 GGAGAGCACAGTGTTCAGGCAACGGAATCCAGAGCCGCGCA-GTCTACGCCCGCCCTCGGCC 240

Qy 340 CACCGACCGCTGGCCCTGCGGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGCC 399

Db 241 CACCGACCGCTGGCCCTGCGGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGCC 300

Qy 400 CACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCCCACCATCTCGCTGTTCAGACGG 459

Db 301 CACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCCCACCATCTCGCTGTTCAGACGG 360

Qy 460 GGAGGAGCCCCCACCCTACAGGGGGGGGGTGCACCCCTCCAGTTCGGGACCCCGAGCAGCA 519

Db 361 GGAGGAGCCCCCACCCTACAGGGGGGGGGTGCACCCCTCCAGTTCGGGACCCCGAGCAGCA 420

Qy 520 GCTGGAACCTGAACCGGAGTTCGGTTCGGCGCACCCCAACAGAACCATCTTCGACAGTGA 579

Db 421 GCTGGAACCTGAACCGGAGTTCGGTTCGGCGCACCCCAACAGAACCATCTTCGACAGTGA 480

Qy 580 CCTGATGGATAGTGCAGGCTGGGGGGGGGGTGCACCCCTGCCGCCCCAGCAGTAACCTCGGGCATCAG 639

Db 481 CCTGATGGATAGTGCAGGCTGGGGGGGGGGTGCACCCCTGCCGCCCCAGCAGTAACCTCGGGCATCAG 540

Qy 640 CGCCACGTGTACGGCAGCGGGGGGGGGTGCATGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG 699

Db 541 CGCCACGTGTACGGCAGCGGGGGGGGGTGCATGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG 600

Qy 700 CATCGGCCACTACCGGGGGTCTCTTCCAGCACCCAGCAGAGCAGTGGGGGGGGGGGGGGGG 759

Db 601 CATCGGCCACTACCGGGGGTCTCTTCCAGCACCCAGCAGAGCAGTGGGGGGGGGGGGGGGG 660

Qy 760 GCTGGAGGGGACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTG 819

||||| 661 GCTGGAGGGACCCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCA-SCATCTG 719
||| : |||||
Qy 820 GAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGGC 879
||||| : |||||
Db 720 GAGCAAGAGAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGGC 779
||||| : |||||
Qy 880 TGGGCTCGTAGGTGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
||||| : |||||
Db 780 TGGGCTCGTAGGTGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839
||||| : |||||
Qy 940 AGGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAAT 999
||||| : |||||
Db 840 AGGCGGGGGSGM--SAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAAT 897
||||| : |||||
Qy 1000 ATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 1059
||||| : |||||
Db 898 ATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA 957
||||| : |||||
Qy 1060 AA 1061
|||
Db 958 AA 959
|||

RESULT 4
AL558881/c
LOCUS
DEFINITION AL558881 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ015YF12 3-PRIME, mRNA sequence.
ACCESSION AL558881
VERSION AL558881.2 GI:31283014
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 951)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12903836.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9945.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ015DC06NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ015DC06NP1.

FEATURES
source
1..951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YF12"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 141 a 275 c 318 g 201 t 16 others
ORIGIN
Query Match 79.6%; Score 844.2; DB 9; Length 951;
Best Local Similarity 99.1%; Pred. No. 1.9e-148;
Matches 846; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 160 GGTGGTGATCACGTGCTGCTGAGCCACTACAAAGTGTCTGCACGGTCTTTCATCAGCCG 219
||| : |||||
Db 854 GGATGTGWTACGTGCTGCTGAGCCACTACAAAGTGTCTGCACGGTCTTTCATCAGCCG 795
||||| : |||||
Qy 220 GCACAGCCAGGGGGGAGAGAGAGAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTC 279
||||| : |||||
Db 794 GAAACAGCCAGGGGGGAGAGAGAGAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTC 735
||||| : |||||
Qy 280 GGAGAGCACAGTGTCAAGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCCCTTCAGCC 339
||||| : |||||
Db 734 GGAGAGCACAGTGTCAAGCAACGGAATCCAGAGCGCAGGTCTACGCCCGCCCTTCAGCC 675
||||| : |||||
Qy 340 CACCGACCGCTGGCCCTGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCC 399
||||| : |||||
Db 674 CACCGACCGCTGGCCCTGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGCC 615
||||| : |||||
Qy 400 CACCTATCCGTACCTGCAGCACGAGATCGACCTGCGCCCGCCACCATCTCGCTGTCTCAGACGG 459
||||| : |||||
Db 614 CACCTATCCGTACCTGCAGCACGAGATCGACCTGCGCCCGCCACCTTCTCGCTTTCAGACGG 555
||||| : |||||
Qy 460 GGAGGAGCCCCACCTTACCAGGGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGAGCAGCA 519
||||| : |||||
Db 554 GGAGGAGCCCCACCTTACCAGGGGGCCCTTGACCCCTCCAGCTTCGGGACCCCGAGCAGCA 495
||||| : |||||
Qy 520 GCTGGAACCTGAACCGGAGTGGTGGCGGCACCCCAACAGAAACCATCTTCGACAGTGA 579
||||| : |||||
Db 494 GCTGGAACCTGAACCGGAGTGGTGGCGGCACCCCAACAGAAACCATCTTCGACAGTGA 435
||||| : |||||
Qy 580 CCTGATGGATAGTCCAGGCTGGCGGGCCCTGCCCGCCAGCAGTAACCTCGGGCATCAG 639
||||| : |||||
Db 434 CCTGATGGATAGTCCAGGCTGGCGGGCCCTGCCCGCCAGCAGTAACCTCGGGCATCAG 375
||||| : |||||
Qy 640 CGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGGCGCCGCCACCTACAGCGAGGT 699
||||| : |||||
Db 374 CGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGGCGCCGCCACCTACAGCGAGGT 315
||||| : |||||
Qy 700 CATCGGCCACTACCGGGTCTCTTCCAGCACACACATCGCGCCCTTAGAGAGCGCATCTG 819
||||| : |||||
Db 314 CATCGGCCACTACCGGGTCTCTTCCAGCACACACATCGCGCCCTTAGAGAGCGCATCTG 255
||||| : |||||
Qy 760 GCTGGAGGGGACCCGGTCCACACACACATCGCGCCCTTAGAGAGCGCATCTG 195
||||| : |||||
Db 254 GCTGGAGGGGACCCGGTCCACACACACATCGCGCCCTTAGAGAGCGCATCTG 195
||||| : |||||
Qy 820 GAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGGC 879
||||| : |||||
Db 194 GAGCAAGAGAGAGATAAACAGAAAGGACACCTCTCTAGGGTCCCCAGGGGGCCGGGC 135
||||| : |||||
Qy 880 TGGGGCTCGTAGGTGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
||||| : |||||
Db 134 TGGGGCTCGTAGGTGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 75
||||| : |||||
Qy 940 AGGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAAT 999
||||| : |||||
Db 74 AGGCGGGGGCGCAGCAACGCATCGTGTGGCCCNCCSTCCACCTCCCTGTGTATATAAT 15
||||| : |||||
Qy 1000 ATTTACATGTGATG 1013
||||| : |||||
Db 14 ATTTACATGTGATG 1

RESULT 5
BX362396/c
LOCUS
DEFINITION BX362396 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ014YN15 3-PRIME, mRNA sequence.
ACCESSION BX362396
VERSION BX362396.1 GI:30378625
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 874)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9945.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ014CG08NP1&cluster=9945.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DJ014CG08NP1.
Location/Qualifiers
1..874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ014YN15"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 125 a 257 c 278 g 180 t 34 others
ORIGIN
Query Match 74.9%; Score 794.4; DB 13; Length 874;
Best Local Similarity 94.3%; Pred. No. 4.3e-139;
Matches 824; Conservative 20; Mismatches 27; Indels 3; Gaps 3;
QY 146 ATGATGGTGATGGTGGTGATC-ACGTGCCTGCTGAGCCACTACAAGCTGTCTGCACG 204
Db ATGATGGTGATSGTGGTSGTGATCTMCGTGCCTGCTGAGCCATTACARGCTGTCTGCACG 815
QY 205 GTCTTTCATCAGCCGGCACAGCCAGGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAG 264
Db GTCTTTCATCAGCCSGCACAGCCAGGGGCGGAGGAGAGMAGATGCCCTKTCTCAGAAAG 755
QY 265 ATGCTGTGGCCCTCGGAGAGACAGTGTCAAGCAACGGGAATCCAGAGCCCGCAGGTCTA 324
Db ATGCTGTGGCCCTCGGAGAGACAGTGTCAAGCAACGGGAATCCAGAGCCCGCA-GTYTA 696
QY 325 CGCCCGCCTCGGCCCCACCGACCGCTGGCCGTGCCGCCCTTCGCCCCAGCGGAGCGCTT 384
Db CGCCCGCCTTCGGCCCCACCGACCGCTGGCCGTGCCGCCCTTCGCCCCAGCGGAGCGCTT 636
QY 385 CCACCGCTTCAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCCACCAT 444
Db CCACCGCTTCAGCCCCACCTATCCGTACCTGNAGNAGAGATNGACNTGCCGCCCCACCAT 576
QY 445 CTCGCTGTCAACGGGGAGGAGCCCCCACCCTACCAGGGGCCCCCTGCACCCTCCAGCTTCG 504
Db CTCGCTGTCAACGGGGAGGAGCCCCCACCCTACCAGGGGCCCCCTGCACCCTCMAGTTTCG 516
QY 505 GGACCCCGAGCAGCAGCTGGAAGTGAACCGGGAGTCGGTGGCGGCACCCCAACAGAAC 564
Db GGACCCCGAGMAGCAGCTGGAAGTGAACCGGGAGTCGGTGGCGGCACCCCAACAGAAC 456
QY 565 CATCTTCGACAGTGACCTGATGGATAGTCCAGGCTGGGGGCCCCCTGCCGCCCCAGCAG 624
Db CATCTTCGACAGTGACCTGATGGATAGTCCAGGCTGGGGGCCCCCTGCCGCCCCAGCAG 396
QY 625 TAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGCGCGCC 684
Db TAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGGCGCGCC 336

QY 685 CACCTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAG 744
Db CACCTACAGCGAGGTCAATCGGCCACTACCCGGGGTCTCTCTCCAGCACCAGCAGCAG 276
QY 745 TGGGCCGCCCTCCTTGTGGAGGGGACCCGGTCTCCACCACACACATCGCGCCCTAGA 804
Db TGGGCCGCCCTCCTTGTGGAGGGGACCCGGTCTCCACCACACACATCGCGCCCTAGA 216
QY 805 GACGCAGCCATCTGGAGCAAGAGAGGATAAACAAGAAAGGACACCCCTCTCTAGSGTCC 864
Db GACGCAGCCATCTGGAGCAAGAGAGGATAAACAAGAAAGGACACCCCTCTCTAGSGTCC 156
QY 865 CCAGGGGGCCCGGCTGGGCTGCGTAGGTGAAAGGCGAGAACACTCCCGGCTTCTTAGA 924
Db CCAGGGGGCCCGGCTGGGCTGCGTAGGTGAAAGGCGAGAACACTCCCGGCTTCTTAGA 96
QY 925 AGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACC 984
Db AGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCANN 36
QY 985 -TCCCTGTGTATAAATATTACATGTGATGTCTG 1017
Db GGNNTNNATATAAATATAAATAATGTGAAGAVDR 2

RESULT 6
BU539219
LOCUS
DEFINITION BU539219 945 bp mRNA linear EST 13-SEP-2002
AGENCOURT 10215265 NIH_MGC_107 Homo sapiens cDNA clone
IMAGE:6569922 5', mRNA sequence.
ACCESSION BU539219
VERSION BU539219.1 GI:22849660
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 945)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2757 row: p column: 18
High quality sequence stop: 663.

FEATURES
source
1..945
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6569922"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 194 a 319 c 287 g 144 t 1 others
ORIGIN

Db 421 CTGAACGGGAGTCGGTGGCGCACCCCAACAGAACCATCTTCGACAGTGACCTGATG 480

Qy 587 GATAGTCCAGGCTGGCGGCCCTGCCCCCAGCAGTAACTCGGGCATCAGCGCCACG 646

Db 481 GATAGTCCAGGCTGGCGGCCCTGCCCCCAGCAGTAACTCGGGCATCAGCGCCACG 540

Qy 647 TGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCAATCGC 706

Db 541 TGCTACGGCAGCGCGGGCGTATGGAGGGCGCGCGCCGCCACCTACAGCGAGGTCAATCGC 600

Qy 707 CACTACCGGGGTCTCTCTCCAGCACAGAGCAGTGGCGCGCCCTCTTGTGTGAG 766

Db 601 CACTACCGGGGTCTCTCTCCAGCACAGAGCAGTGGCGCGCCCTCTTGTGTGAG 660

Qy 767 GGGACCGGCTCCACCACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAA 826

Db 661 GGGACCGGCTCCACCACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAA 720

Qy 827 GAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGCTGGG-- 884

Db 721 GAGAGGATAAACAGAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGCTGGGCG 780

Qy 885 -CTGCGTAGGTGAAAGGCGAGACACTCCGCGCTTCTTAGAGAGGAGTGAGAGGAGGC 943

Db 781 TTGCTAAGGAAAGGACACACCTCCCGCGCTTCTTAAAGAGGGTGGAAACGGC 840

Qy 944 GGGGGGGCAGCA 956

Db 841 AGGCGGGGGCA 853

RESULT 8

LOCUS BQ015170/c

DEFINITION UI-H-ED1-axw-k-20-0-UI.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone IMAGE:5834635 3', mRNA sequence.

ACCESSION BQ015170

VERSION BQ015170.1 GI:19740071

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 782)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1..782

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5834635"

/tissue_type="Chondrosarcoma"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI_CGAP_ED1"

/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; NCI_CGAP_ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line CS5. The library was constructed according to Bonaldo

, Lennon and Soares, Genome Research, 6:791-806, 1996.
First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.

TAG LIB=UI-H-ED1
TAG_TISSUE=chondrosarcoma
TAG_SEQ=CGTCAAGGCT"
BASE COUNT 109 a 223 c 271 g 176 t 3 others
ORIGIN

Query Match 72.4%; Score 768.4; DB 12; Length 782;
Best Local Similarity 98.8%; Pred. No. 3.2e-134;
Matches 772; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 281 GAGAGCACAGTGTCCAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCC 340

Db 782 GAGAGCACAGTGTCCAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCGGCC 723

Qy 341 ACCGACCGCTGGCGCTGCCGCCCTTCGCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCC 400

Db 722 ACCGACCGCTGGCGCTGCCGCCCTTCGCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCC 663

Qy 401 ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCACCACCATCTCGTGTTCAGACGGG 460

Db 662 ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCCACCACCATCTCGTGTTCAGACGGG 603

Qy 461 GAGGAGCCCCCACCTTACCAGGGCCCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGCAG 520

Db 602 GAGGAGCCCCCACCTTACCAGGGCCCCCTGCACCCCTCCAGTTCGGGACCCCGAGCAGCAG 543

Qy 521 CTGGAACCTGAACCGGGAGTCGGTGCAGCGCACCCCGCCCAACAGAACCATCTTTCGACAGTGAC 580

Db 542 CTGGAACCTGAACCGGGAGTCGGTGCAGCGCACCCCGCCCAACAGAACCATCTTTCGACAGTGAC 483

Qy 581 CTGATGGATAGTGCCAGGCTGGCGGGCCCCCTGCCCCCAGCAGTAACCTCGGGCATCAGC 640

Db 482 CTGATGGATAGTGCCAGGCTGGCGGGCCCCCTGCCCCCAGCAGTAACCTCGGGCATCAGC 423

Qy 641 GCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGGGCGCGCCGCCACCTACAGCGAGGTC 700

Db 422 GCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGGGCGCGCCGCCACCTACAGCGAGGTC 363

Qy 701 ATCGGCCACTACCCGGGGTCTCTTCCAGCACCGAGCAGTGGGGCGCCCTCCTTG 760

Db 362 ATCGGCCACTACCCGGGGTCTCTTCCAGCACCGAGCAGTGGGGCGCCCTCCTTG 303

Qy 761 CTGGAGGGGACCCGGTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 820

Db 302 CTGGAGGGGACCCGGTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGG 243

Qy 821 AGCAAAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGCGGCT 880

Db 242 AGCAAAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCAGGGGGCGCGGCT 183

Qy 881 GGGGCTCGTAGGTGAAAAGGCAGAACACTCCCGCGCTTCTTAGAGAGGAGTGAAGAGAA 940

Db 182 GGGGTTGCTAGGTGAAAAGGCAGAACACTCCCGCGCTTCTTAGAGAGGAGTGAAGAGAA 123

Qy 941 GCGGGGGGGCGCAGCAACCGCATCGTGTGGGCCCTCCCTCCCAACCTCCCTGTGTATAATA 1000

Db 122 GCGGGGGGGCGCAGCAACCGCATCGTGTGGGCCCTCCCTCCCAACCTCCCTGTGTATAATA 63

Qy 1001 TTTACATGTGATGTCTGGTCTGAATGCACAAAGCTAAGAGAGCTTGCAAAAAA 1060

Db 62 TTTACATGTGATGTCTGGTCTGAATGCACAAAGTTAAGAGAGCTTGCAAAAAA 3

Qy 1061 A 1061


```
Db      2 A 2
RESULT 9
BQ575741/c
LOCUS
DEFINITION
  UI-H-EZ1-bbg-h-14-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone
  UI-H-EZ1-bbg-h-14-0-UI 3', mRNA sequence.
ACCESSION
  BQ575741
VERSION
  BQ575741.1 GI:21479058
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 729)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
JOURNAL
  Unpublished
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of
  Orthopedics
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be obtained
  from Dr. M. Bento Soares, bento-soares@uiowa.edu
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  Location/Qualifiers
    1..729
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-H-EZ1-bbg-h-14-0-UI"
      /tissue_type="Chondrosarcoma Grade II"
      /dev_stage="Adult"
      /lab_host="DH10B (Life Technologies)"
      /clone_lib="NCI_CGAP_Ch2"
      /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
      with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
      NCI CGAP Ch2 is a normalized cDNA library containing the
      following tissue(s): Chondrosarcoma Grade II. The library
      was constructed according to Bonaldo, Lennon and Soares,
      Genome Research, 6:791-806, 1996. First strand cDNA
      synthesis was primed with an oligo-dT primer containing a
      Not I site. Double stranded cDNA was ligated to an EcoR I
      adaptor, digested with Not I, and cloned directionally
      into pT7T3-Pac vector. The oligonucleotide used to prime
      the synthesis of first-strand cDNA contains a library tag
      sequence that is located between the Not I site and the
      (dT)18 tail. The sequence tag for this library is
      TGATCAGCT.
      TAG_LIB=UI-H-EZ1
      TAG_TISSUE=grade-2-chondrosarcoma
      TAG_SEQ=ATCTAATATG"
BASE COUNT 101 a 213 c 253 g 162 t
ORIGIN
  Query Match 68.3%; Score 724.4; DB 13; Length 729;
  Best Local Similarity 99.9%; Pred. No. 5.8e-126;
  Matches 725; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 336 GGGCCACGACCGCCTGGCCGTGCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCC 395
  |||||||
Db 729 GGGCCACGACCGCCTGGCCGTGCGCCCTTCGCCAGCGGGAGCGCTTCCACCGCTTCC 670
  |||||||
Qy 396 AGCCCACTATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGTGTGAC 455
  |||||||
Db 669 AGCCCACTATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGTGTGAC 610
  |||||||
```

```
Qy 456 ACGGGGAGGAGCCCCACCTTACCAGGGCCCCCTGCACCTTCCAGCTTCGGGACCCGAGC 515
  |||||||
Db 609 ACGGGGAGGAGCCCCACCTTACCAGGGCCCCCTGCACCTTCCAGCTTCGGGACCCGAGC 550
  |||||||
Qy 516 AGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCGACCCCAACAGAACCATCTTCGACA 575
  |||||||
Db 549 AGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCGACCCCAACAGAACCATCTTCGACA 490
  |||||||
Qy 576 GTGACCTGATGGATAGTCCAGGCTGGGGCGGCCCTTGCCTCCCGCCAGCAGTAACCTCGGGCA 635
  |||||||
Db 489 GTGACCTGATGGATAGTCCAGGCTGGGGCGGCCCTTGCCTCCCGCCAGCAGTAACCTCGGGCA 430
  |||||||
Qy 636 TCAGCGCCACGTGTACGGCAGCGGGCGGGCGGCATGGAGGGGGCGCCGCCACCTACAGCG 695
  |||||||
Db 429 TCAGCGCCACGTGTACGGCAGCGGGCGGGCGGCATGGAGGGGGCGCCGCCACCTACAGCG 370
  |||||||
Qy 696 AGGTTCATCGGCCACTACCGGGGGTCTCTTCCAGCAGCAGCAGTGGGGCGCCCT 755
  |||||||
Db 369 AGGTTCATCGGCCACTACCGGGGGTCTCTTCCAGCAGCAGCAGTGGGGCGCCCT 310
  |||||||
Qy 756 CTTGTCTGGAGGGGACCGGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCA 815
  |||||||
Db 309 CTTGTCTGGAGGGGACCGGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCA 250
  |||||||
Qy 816 TCTGGACCAAGAGAGGATAAAGAGGACACCCCTCTCTAGGGTCCCCAGGGGGGCC 875
  |||||||
Db 249 TCTGGACCAAGAGAGGATAAAGAGGACACCCCTCTCTAGGGTCCCCAGGGGGGCC 190
  |||||||
Qy 876 GGGCTGGGGCTCGTAGTGAAGGAGGAGAACTCCGCGCTTCTTAGAAGAGGAGTAGAG 935
  |||||||
Db 189 GGGCTGGGGCTCGTAGTGAAGGAGGAGAACTCCGCGCTTCTTAGAAGAGGAGTAGAG 130
  |||||||
Qy 936 AGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTAT 995
  |||||||
Db 129 AGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTAT 70
  |||||||
Qy 996 AAATATTACATGTGATGTCGTGTTGAATGCACAAAGTAAAGAGAGCTTGCAAAAAAAA 1055
  |||||||
Db 69 AAATATTACATGTGATGTCGTGTTGAATGCACAAAGTAAAGAGAGCTTGCAAAAAAAA 10
  |||||||
Qy 1056 AAAAAA 1061
  |||||
Db 9 AAGAAA 4
  |||||
RESULT 10
BM677602/c
LOCUS
DEFINITION
  UI-E-E01-aid-p-09-0-UI.s1 UI-E-E01 Homo sapiens cDNA clone
  UI-E-E01-aid-p-09-0-UI 3', mRNA sequence.
ACCESSION
  BM677602
VERSION
  BM677602.1 GI:18987498
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 730)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
JOURNAL
  97044477
MEDLINE
  8889548
PUBMED
  8889548
COMMENT
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA.
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  Tissue Procurement: Dr. Gregg Hageman
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
```

cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1. .730
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E01-aid-p-09-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E01"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG_LIB=UI-E-E01
TAG_TISSUE=human fetal eye
TAG_SEQ=CGCGTATACC"
BASE COUNT 103 a 210 c 251 g 164 t 2 others
ORIGIN

Query Match 67.9%; Score 720.2; DB 12; Length 730;
Best Local Similarity 99.3%; Pred. No. 3.5e-125;
Matches 722; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 335 CGGCCCCACGACCGCTGGCCGTGCGCCGCTTTCGCCACGCGGAGCGCTTCCACCGCTTC 394
Db |||||||
QY 730 CGGCCCCACGACCGCTTGGCCGTGCGCCGCTTTCGCCACGCGGAGCGCTTCCACCGCTTC 671
Db |||||||
QY 395 CAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCGGCCACCATCTCGTGTCA 454
Db |||||||
QY 670 CAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACTTGGCACCCACCATCTCGTGTCA 611
Db |||||||
QY 455 GACGGGGAGGAGCCCCACCTACGAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAG 514
Db |||||||
QY 610 GACGGGGAGGAGCCCCACCTACGAGGGCCCCCTGCACCTCCAGCTTCGNGACCCCGAG 551
Db |||||||
QY 515 CAGCAGCTGGAACCTGAACCGGGAGTCGGTGCGCGCACCCCCAACAGAACCATCTTCGAC 574
Db |||||||
QY 550 CAGCAGCTGGAACCTGAACCGGGAGTCGGTGCGCGCACCCCCAACAGAACCATCTTCGAC 491
Db |||||||
QY 575 AGTGACCTGATGATAGTGCCAGGCTGGCGGCGCCCTGCCCCCCCCCAGCAGTAACTCGGGC 634
Db |||||||
QY 490 AGTGACCTGATGATAGTGCCAGGCTGGCGGCGCCCTGCCCCCCCCCAGCAGTAACTCGGGC 431
Db |||||||
QY 635 ATCAGCGCCACGTGCTACGGCAGCGGGCGCGCATGGAGGGCGCGCCGCCACCTACAGC 694
Db |||||||
QY 430 ATCAGCGCCACGTGNTACGGCAGCGGGCGCGCATGGAGGGCGCGCCGCCACCTACAGC 371
Db |||||||
QY 695 GAGGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACGACGAGCAGTGGGGCCGCC 754
Db |||||||
QY 370 GAGGTTCATCGGCCACTACCCGGGGTCTCTCTCCAGCACGACGAGCAGTGGGGCCGCC 311
Db |||||||
QY 755 TCCTTGCTGGAGGGGACCCGGCTCCACCACACACATATCGGCCCCCTAGAGAGCGGAGCC 814
Db |||||||
QY 310 TCCTTGCTGGAGGGGACCCGGCTCCACCACACACATATCGGCCCCCTAGAGAGCGGAGCC 251
Db |||||||

QY 815 ATCTGGAGCAAAAGAGAGGATAAAACAGAAAGGACACCTCTCTAGGTTCCCCAGGGGGC 874
Db |||||||
QY 250 ATCTGGAGCAAAAGAGAGGATAAAACAGAAAGGACACCTCTCTAGGTTCCCCAGGGGGC 191
Db |||||||
QY 875 CGGGCTGGGGCTGCGTAGGTGAAAAGGCGAGAACACTCCGCGCTTCTTAGAAGAGGAGTGA 934
Db |||||||
QY 190 CGGGCTGGGGCTGCGTAGGTGAAAAGGCGAGAACACTCCGCGCTTCTTAGAAGAGGAGTGA 131
Db |||||||
QY 935 GAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTA 994
Db |||||||
QY 130 GAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTA 71
Db |||||||
QY 995 TAAATATTTACATGTGATGTCTGTGATGCAATGCACAAAGCTTAAGAGAGCTTGCAAAAAA 1054
Db |||||||
QY 70 TAAATATTTACATGTGATGTCTGTGATGCAATGCACAAAGCTTAAGAGAGCTTGCAAAAAA 11
Db |||||||
QY 1055 AAAAAA 1061
Db 10 AAAAAA 4

RESULT 11
BU683523/c
LOCUS BU683523 728 bp mRNA linear EST 07-OCT-2002
DEFINITION UI-CF-EC1-acg-e-09-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
UI-CF-EC1-acg-e-09-0-UI 3', mRNA sequence.
ACCESSION BU683523
VERSION BU683523.1 GI:23535533
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 728)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
1. .728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acg-e-09-0-UI"
/tissue_type="Lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AAGTGCTTAC.

TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGTGCTTAC"

BASE COUNT	103 a	209 c	252 g	163 t	1 others
ORIGIN					

Query Match 66.9%; Score 709.8; DB 13; Length 728;
Best Local Similarity 99.4%; Pred. No. 3.1e-123;
Matches 722; Conservative 0; Mismatches 3; Indels 1;

QY.	336	GGCCACCGACCGCTGGCCGTCGTGCCGCCCTTGCGCCAGCGGAGCGCTTCCACCGCTTCC	395
Db	728	GGCCACCGACCGCTGGCCGTCGTGCCGCCCTTGCGCCAGCGGAGCGCTTCCACCGCTTCC	669
QY	396	AGCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCACTCTCGCTGTCTCAG	455
Db	668	AGCCACCTATCCGTACCTGCAGCATGAGATCGACCTGCCGCCACCACTCTCGCTGTCTCAG	609
QY	456	ACGGGAGGAGCCCCACCCCTACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGC	515
Db	608	ACGGNGAGGAGCCCCACCCCTACCAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGC	549
QY	516	AGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCACTCTTCGACA	575
Db	548	AGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGGCACCCCCAAACAGAACCACTCTTCGACA	489
QY	576	GTGACCTGATGATAGTGCCAGGCTGGCGGCCCTGCTGCCCCCCCAGCAGTAACTCGGGCA	635
Db	488	GTGACCTGATGATAGTGCCAGGCTGGCGGCCCTGCTGCCCCCCCAGCAGTAACTCGGGCA	429
QY	636	TCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCCGCCACCTACAGCG	695
Db	428	TCAGCGCCACGTGCTACGTACGCGGGCGGCATGGAGGGCGCCGCCACCTACAGCG	369
QY	696	AGGTTCATCGGCCACTACCCGGGTCTCTCTTCAGCACCCAGCAGAGCAGTGGSCCGCCCT	755
Db	368	AGGTTCATCGGCCACTACCCGGGTCTCTCTTCAGCACCCAGCAGAGCAGTGGSCCGCCCT	309
QY	756	CCTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCA	815
Db	308	CCTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCCCTAGAGAGCGCAGCCA	249
QY	816	TCTGAGCAAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCCAGGGGGGCC	875
Db	248	TCTGAGCAAGAGAAGGATAAACAGAAAGGACACCCCTCTCTAGGGTCCCCAGGGGGGCC	189
QY	876	GGCTGGGGCTGCGTAGGTGAAAAGCGAGAACACTCCCGCGCTTCTTAGAAGAGAGTGAG	935
Db	188	GGCTGGGGCTGCGTAGGTGAAAAGCGAGAACACTCCCGCGCTTCTTAGAAGAGAGTGAG	129
QY	936	AGGAAGCGGGGGCGCAGCAACGCATCGTGTGSCCCTCCCTCCACCTCCCTGTGTAT	995
Db	128	AGGAAGCGGGGGCGCAGCAACGCATCGTGTGSCCCTCCCTCCACCTCCCTGTGTAT	70
QY	996	AAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAAAA	1055
Db	69	AAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAAAA	10
QY	1056	AAAAAA 1061	
Db	9	AAAAAA 4	

RESULT 12	ACCESSION	REFERENCE
BUG02918	VERSION	AUTHORS
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	JOURNAL
	ORGANISM	COMMENT

CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6497853"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC 142"
/notes="Vector: pDNR-LIB; Site 1: SfiI
Site 2: SfiI (ggcgctcgcc); Double-
prepared from a pool of 40 cell line p
- 2%, blood - 33.4%, brain - 5.6%, bre
4%, connective tissue - 1.4%, eye - 1%
kidney - 2.2%, liver - 5.7%, lung - 10
ovary - 4%, pharynx - 2.5%, prostate -
gland - 1.3%, and skin - 2.3%). 5' and
used in cloning as follows:
5'-AACGAGTGGTATCAACGACGAGTGGCATTACGGC
5'-ATTCTAGAGCGCGGCGCCGACATG-dT(30)NNNN
enriched library was constructed using
Creator SMART kit and size-selected to
size fraction (other fractions present
Library created in the laboratory of M
NIH). Note: this is a NIH MGC Library.

```

	NMN / NOCE: CH18 IS a NMN		
	173 a	303 c	251 g
BASE COUNT	123 t		
ORIGIN			

Query Match 65.8%; Score 697.8; DB 13; Length 850;
Best Local Similarity 92.9%; Pred. No. 5.4e-121;
Matches 787; Conservative 0; Mismatches 52; Indels 8; Gaps 5;

QY	142	GGTGATGATGGTGTGGTGGTGTGATCACGTGCCCTGCTAGCCACTACAAGCTGTCTGC	201
D _b	1	GGTGATGATGGTGTGGTGGTGTGATCACGTGCCCTGCTAGCCACTACAAGCTGTCTGC	60
QY	202	ACGGTCCTTCATCAGCCGSCACAGCCAGGGCGGAGGAGAAGATGCCCTGTCTCTCAGA	261
D _b	61	ACGGTCCTTCATCAGCCGSCACAGCCAGGGCGGAGGAGAAGATGCCCTGTCTCTCAGA	120
QY	262	AGGATGCCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGCAGGT	321
D _b	121	AGGATGCCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGCAGGT	180
QY	322	CTACGCCCCGCGCTCGGCCAACCGACCCTGGCCGTGCCGCCCTTCGCCCCAGCGGAGCG	381

Db 181 CTACGCCCGCCTCGGCCACCGACCGCCTGGCCGTGCGCCTTCGCCCGAGCGGACCG 240
QY 382 CTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCAC 441
Db 241 CTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCAGAGATCGACCTGCCGCCAC 300
QY 442 CATCTCGCTGTGACACGGGAGGAGCCCCCACCCTACCAGGSCCCTGCACCCCTCCAGCT 501
Db 301 CATCTCGCTGTGACACGGGAGGAGCCCCCACCCTACCAGGSCCCTGCACCCCTCCAGCT 360
QY 502 TCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCATGGAGGGCGCC 681
Db 361 TCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCATGGAGGGCGCC 420
QY 562 AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGTGGGCGGCCCTGCCGCCCCAG 621
Db 421 AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGTGGGCGGCCCTGCCGCCCCAG 480
QY 622 CAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGCGCC 681
Db 481 CAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGGCATGGAGGGCGCC 540
QY 682 GCCACCTACAGCGAGGTGATCGGCCACTACCGGGGTCTCTCTCCAGCACCAGCAGAG 741
Db 541 GCCACCTACAGCGAGGTGATCGGCCACTACCGGGGTCTCTCTCCAGCACCAGCAGAG 600
QY 742 CAGTGGGCGG-CCCTCCTTGTGG-AGGGGACCGCGTCCACACACACATCGCGGCC 799
Db 601 CAGTGGGCGGCCCTCCTTGTGGAAGGGGACCCGGTTCACACACACAAATCCGGGCC 660
QY 800 CTAGAGAGCGCAGCCA-TCTGAGCAAAAGAGAGGATAAAGAGGAGGACACCTCTCTA 858
Db 661 CTAAGAGCGCAGCCATTCTGGAAGCAAGCAGATGATAAAGAGGAGGACCCCTCTCTA 720
QY 859 GGGTCCCAGGGGGCC--GGGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACT---CCG 913
Db 721 GGGTCCCAGGGGGCCCGGGCTGGGGCTGTCTAGGTGAAAAGGAAACACTTCCCG 780
QY 914 CGTTCTTAGAAGAGGAGTGAGAGGAAGGGGGGGCGGCGCAGCAACGATCGTGTGGCCT 973
Db 781 CTTCTTAAAAACGATGTATAAGAAAGCCGGGGGGCCAAACCCCTTTGGGTGGGCC 840
QY 974 CCCCTCC 980
Db 841 CTCCTCC 847

RESULT 13
BQ690750
LOCUS
DEFINITION BQ690750 890 bp mRNA linear EST 15-JUL-2002
5', mRNA sequence.
ACCESSION BQ690750.1 GI:21816066
VERSION BQ690750
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 890)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM2368 row: h column: 14

FEATURES
source
High quality sequence stop: 627.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6209341"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 168 a 303 c 273 g 146 t
ORIGIN

Query Match 64.1%; Score 680.2; DB 13; Length 890;
Best Local Similarity 95.9%; Pred. No. 1.le-117;
Matches 709; Conservative 0; Mismatches 28; Indels 2; Gaps 1;
QY 1 TCCTCCTTGGGTTTCGGGTGAAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 60
Db 24 TCCTCCTTGGGTTTCGGGTGAAAGCGCTTGGGGTTTCAGTGGGCCATGATCCCCGAGCTGC 83
QY 61 TGGAGAACTGAAGCGGCGGCTCTCCTGCGAAACCCAGGCAATGGCGGAGCTGGAGTTGT 120
Db 84 TGGAGAACTGAAGCGGCGGAGTCTCCTGCGAAACCCAGGCAATGGCGGAGCTGGAGTTGT 143
QY 121 TCAGATCATCATCATCTGTTGGTGGTATGATGGTATGGTGGTGGTGGTGGTGGTGGTGGT 180
Db 144 TCAGATCATCATCATCTGTTGGTGGTATGATGGTATGGTGGTGGTGGTGGTGGTGGTGGT 203
QY 181 GAGCCACTACAAGCTGTCTGACCGTCTCTTCATCAGCCGACAGCCAGGGCGGAGGAG 240
Db 204 GAGCCACTACAAGCTGTCTGACCGTCTCTTCATCAGCCGACAGCCAGGGCGGAGGAG 263
QY 241 AGAAGATGCCCTGTCTCCTCAGAAGGATGCCCTGTGGCCCTCGGAGAGCACAGTGTCAAGGCAA 300
Db 264 AGAAGATGCCCTGTCTCCTCAGAAGGATGCCCTGTGGCCCTCGGAGAGCACAGTGTCAAGGCAA 323
QY 301 CGGAATCCCAGAGCGCGCAGGTCTACGCCCCCGCTCGGCCCGCTCGGCCCGCTGGCCGTGCC 360
Db 324 CGGAATCCCAGAGCGCGCAGGTCTACGCCCCCGCTCGGCCCGCTCGGCCCGCTGGCCGTGCC 383
QY 361 GCCCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCA 420
Db 384 GCCCTTCGCCCGAGCGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACCTGCAGCA 443
QY 421 CGAGATCGACCTGCCCGCCACCATCTCGCTGTGACAGCGGAGGAGCCCCACCTACCA 480
Db 444 CGAGATCGACCTGCCCGCCACCATCTCGCTGTGACAGCGGAGGAGCCCCACCTACCA 503
QY 481 GGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTC 540
Db 504 GGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTC 563
QY 541 GGTGCGCGCACCCCCCAACAGAACCATCTTCGACAGTGAACCTGTGATGATGATGATGATGATGAT 600
Db 564 GGTGCGCGCACCCCCCAACAGAACCATCTTCGACAGTGAACCTGTGATGATGATGATGATGATGAT 623
QY 601 GGGCGGCCCTTGCCTCCCGCCAGCAGTAACTCGGGCATCAGCGCCACGTGTACGGCAGCGG 660
Db 624 GGGCGGCCCTTGCCTCCCGCCAGCAGTAACTCGGGCATCAGCGCCACGTGTACAGCAGCGG 683
QY 661 CGGGCGCATGGA--GGGGCGGCCGCCACCTACAGCGAGGTGATCGGGCCACTACCCCGGG 718
Db 684 CGGGCGCATGGAAGGGGGCCCCCGCCACCTACAGGCAAGGCATCTGGCCACCTACCCCG 743

QY 719 TCCTCCTTCCAGCACCAGC 737 655 bp mRNA linear EST 15-JUL-2002
Db 744 GGGTCCTCCTTTTCCAGC 762 5', mRNA sequence.

RESULT 14
BQ691705
LOCUS BQ691705 655 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8046876 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6208939
ACCESSION BQ691705
VERSION BQ691705.1 GI:21817021
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 655)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2367 row: g column: 20
High quality sequence stop: 645.
Location/Qualifiers
1..655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6208939"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 155 a 216 c 193 g 91 t

Query Match 61.5%; Score 652; DB 13; Length 655;
Best Local Similarity 100.0%; Pred. No. 2.2e-112;
Matches 652; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 TACCTGCAGCAGATCGACCTGCCGCCACCACCATCTCGTGTCTCAGACGGGGAGGAGCCC 469
Db 1 TACCTGCAGCAGATCGACCTGCCGCCACCACCATCTCGTGTCTCAGACGGGGAGGAGCCC 60

QY 470 CCACCTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACGTG 529
Db 61 CCACCTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGGAACGTG 120

QY 530 AACCGGAGTCGGTGCAGCCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGAT 589
Db 121 AACCGGAGTCGGTGCAGCCACCCCAACAGAACCATCTTCGACAGTGACCTGATGGAT 180

QY 590 AGTGCCAGGCTGGGGGGCCCCCTGCCGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACGTGC 649
Db 181 AGTGCCAGGCTGGGGGGCCCCCTGCCGCCCGCCAGCAGTAACCTCGGGCATCAGCGCCACGTGC 240

QY 650 TACGGCAGCGGGGGCGCATGGAGGGGGCGCCGCCACCTACAGCGAGGTTCATCGGGCCAC 709

Db 241 TAGGGCAGCGGGGGCGCATGGAGGGCGCCGCCACCTACAGCGAGGTTCATCGGCCAC 300
QY 710 TACCCGGGGTCTCTTCCAGCACCAGCAGCAGTGGGGCGCCCTCTTGTGCTGAGGGG 769
Db 301 TACCCGGGGTCTCTTCCAGCACCAGCAGCAGTGGGGCGCCCTCTTGTGCTGAGGGG 360

QY 770 ACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAG 829
Db 361 ACCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAG 420

QY 830 AAGGATAAACAGAAAGGACACCTCTCTAGGCTCCCGGGGGCGGGCTGGGCTGCG 889
Db 421 AAGGATAAACAGAAAGGACACCTCTCTAGGCTCCCGGGGGCGGGCTGGGCTGCG 480

QY 890 TAGGTGAAAAGGCAGAACACTCCCGCTTCTTAGAAGAGAGTGAGAGGAAGCGGGGGG 949
Db 481 TAGGTGAAAAGGCAGAACACTCCCGCTTCTTAGAAGAGAGTGAGAGGAAGCGGGGGG 540

QY 950 CGCAGCAACGCATCGTGGGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1009
Db 541 CGCAGCAACGCATCGTGGGGCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 600

QY 1010 GATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAA 1061
Db 601 GATGTCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCACAAAAA 652

RESULT 15
AI761441/c
LOCUS AI761441 693 bp mRNA linear EST 20-DEC-1999
DEFINITION wg55f07.x1 Soares NSF F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone IMAGE:2369989 3', similar to TR:O15168 O15168 CLONE 22. [3]
TR:O15167 TR:O15165 ; mRNA sequence.

ACCESSION AI761441
VERSION AI761441.1 GI:5177108
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 693)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1105 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 464.
Location/Qualifiers
1..693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2369989"
/lab_host="DH10B"
/clone_lib="Soares NSF F8_9W_OT_PA_P_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NbHSF pool 1: 309384-310919, 323208-325895 Soares Nb2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares Nb2HF8-9W pool 1.

FEATURES
source

758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NbHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 101 a 205 c 237 g 149 t 1 others
ORIGIN

Query Match 60.7%; Score 644.2; DB 9; Length 693;
Best Local Similarity 98.1%; Pred. No. 6.2e-111;
Matches 683; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY	359	CCGCCCTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACACCTATCCGTACCTGCAG	418
Db	693	CGGCTTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCCAGCCACCTATCCGTACNTGCAG	634
QY	419	CACGAGATCGACCTGCCGCCACCATCTCGCTGTCTCAGACGGGAGGAGCCCCACCCCTAC	478
Db	633	CACGAGATCGA-CTGCCGCCACCATCTCGCTGTCTCAGACGGG--AGGAGCCCCACCCCTAC	577
QY	479	CAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGAGCTGGAACCTGAACCCGGGAG	538
Db	576	CAGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGAGCTGGAACCTGAACCCGGGAG	517
QY	539	TCGCTGCGCGCACCCCCAAACAGAACCATCTTCGACAGTGCCTGATGGATAGTCCAGG	598
Db	516	TCGCTGCGCGCACCCCCAAACAGAACCATCTTCGACAGTGCCTGATGGATAGTCCAGG	457
QY	599	CTGGCGGCGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGGCCCCACGTCGTACGGCAGC	658
Db	456	CTGGCGGCGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGGCCCCACGTCGTACGGCAGC	397
QY	659	GGCGGGCGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCCGGGG	718
Db	396	GGCGGGCGCATGGAGGGGCGCGCCGCCACCTACAGCGAGGTCTATCGGCCACTACCCCGGGG	337
QY	719	TCCTCCTTCCAGCACAGCAGCAGTGGGCGCCCTCTCTGTCTGGA-GGGGACCCGGCT	777
Db	336	TCCTCCTTCCAGCACAGCAGCAGTGGGCGCCCTCTCTGTCTGGAAGGGGGACCCGGGTT	277
QY	778	CCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAAGAGAGGATAA	837
Db	276	CCCCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCAAAAGAGAGGATAA	217
QY	838	ACAGAAAGGACACCCCTCTCTAGGGTCCCCAGGGGGCGCGGCTGGGGCTGCGTAGGTGAA	897
Db	216	ACAGAAAGGACACCCCTCTCTAGGGTCCCCAGGGGGCGCGGCTGCGTAGGTGAA	157
QY	898	AAGGCAGAACACTCCCGCGCTTTTAGAAGAGGAGTGAAGGAAGCGGGGGCGCAGCAA	957
Db	156	AAGGCAGAACACTCCCGCGCTTTTAGAAGAGGAGTGAAGGAAGCGGGGGCGCAGCAA	97
QY	958	CGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTCTG	1017
Db	96	CGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGTCTG	37
QY	1018	GTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1053
Db	36	GTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA	1

Search completed: December 8, 2003, 03:56:09
Job time : 2647 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 01:56:10 ; Search time 1262 Seconds
(without alignments)
2794.246 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tcctccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1034.4	97.5	1141	13 US-10-301-822-208	Sequence 208, App
2	1034.4	97.5	1141	15 US-10-205-823-412	Sequence 412, App
3	1034.4	97.5	1850	13 US-10-241-220-44	Sequence 44, Appl
4	1034.4	97.5	4527	13 US-09-821-812-2	Sequence 2, Appli
5	1022.4	96.4	1140	13 US-10-390-045-1	Sequence 1, Appli
6	1021.2	96.2	1066	14 US-10-098-841-71	Sequence 71, Appl
7	961.8	90.7	969	11 US-09-796-753-55	Sequence 55, Appl
8	956	90.1	4839	13 US-10-241-220-119	Sequence 119, App
9	956	90.1	4839	13 US-10-269-909-84	Sequence 84, Appl
10	956	90.1	4839	13 US-10-269-909-85	Sequence 85, Appl
11	800.6	75.5	1321	10 US-09-934-249-1	Sequence 1, Appli
12	795.2	74.9	806	13 US-10-241-220-45	Sequence 45, Appl
13	757.4	71.4	759	13 US-10-390-045-2	Sequence 2, Appli
14	752.6	70.9	861	10 US-09-934-249-3	Sequence 3, Appli
15	728.8	68.7	1583	15 US-10-000-256A-32	Sequence 32, Appl

C	16	644.2	60.7	693	10	US-09-934-249-14	Sequence 14, Appl
	17	523.8	49.4	878	10	US-09-934-249-12	Sequence 12, Appl
	18	521.4	49.1	1713	11	US-09-796-753-57	Sequence 57, Appl
	19	350.8	33.1	2170	13	US-10-094-749-197	Sequence 197, App
	20	350	33.0	8093	10	US-09-934-249-16	Sequence 16, Appl
	21	349	32.9	921	13	US-10-251-598-7	Sequence 0, Appli
	22	349	32.9	8065	13	US-10-251-598-6	Sequence 0, Appli
	23	332.8	31.4	467	11	US-09-918-995-2074	Sequence 2074, Ap
	24	281	26.5	867	13	US-10-251-598-8	Sequence 0, Appli
	25	249.2	23.5	475	10	US-09-934-249-15	Sequence 15, Appl
	26	218.2	20.6	368	10	US-09-783-590-3464	Sequence 3464, Ap
	27	60	5.7	60	13	US-09-908-975-13620	Sequence 13620, A
	28	59.2	5.6	65	10	US-09-783-590-3488	Sequence 3488, Ap
C	29	54.8	5.2	5452	13	US-10-017-161-1481	Sequence 1481, Ap
C	30	54.6	5.1	3133	13	US-10-017-161-1483	Sequence 1483, Ap
C	31	53	5.0	1117	13	US-10-017-161-1403	Sequence 1403, Ap
	32	50.2	4.7	2316	15	US-10-156-761-4416	Sequence 4416, Ap
C	33	50.2	4.7	9025608	15	US-10-156-761-1	Sequence 1, Appli
	34	49	4.6	15738	13	US-10-329-079-46	Sequence 46, Appl
	35	49	4.6	61944	13	US-10-329-079-34	Sequence 34, Appl
C	36	48.4	4.6	11304	11	US-09-764-891-5801	Sequence 5801, Ap
	37	48	4.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
C	38	47.4	4.5	1215	13	US-09-953-348-42	Sequence 42, Appl
C	39	47.4	4.5	1215	15	US-10-267-255-42	Sequence 42, Appl
	40	47.4	4.5	53500	13	US-09-953-348-76	Sequence 76, Appl
	41	47.4	4.5	53500	15	US-10-267-255-76	Sequence 76, Appl
	42	47.2	4.4	1233	15	US-10-166-087-9	Sequence 9, Appli
C	43	47.2	4.4	32539	15	US-10-166-087-1	Sequence 1, Appli
	44	46.8	4.4	936	15	US-10-156-761-2133	Sequence 2133, Ap
	45	46.8	4.4	2715	15	US-10-156-761-2245	Sequence 2245, Ap

ALIGNMENTS

RESULT 1

US-10-301-822-208
; Sequence 208, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(854)
US-10-301-822-208

Query Match 97.5%; Score 1034.4; DB 13; Length 1141;
Best Local Similarity 99.2%; Pred. No. 8.6e-274;

Matches 1050; Conservative		0; Mismatches	6; Indels	2; Gaps	1;
Qy	4	TCCTTGGGTT	CGGGTGAAGCGCTTGGGGTT	CAGTGGCCCATGATCCCGAGCTGCTGG	63
Db	1	TCCTTGGGTT	CGGGTGAAGCGCCTGGGGTTC	-GTGGCCATGATCCCGAGCTGCTGG	58
Qy	64	AGAACTGAAG	CGGACGGTCTCTCTGCGAAACCAGGCAATGGCGGAGCTGGAGTTGTTCA	123	
Db	59	AGAACTGAAG	CGGACAGTCTCTCTGCGAAACCAGGCAATGGCGGAGCTGGAGTTGTTCA	118	
Qy	124	GATCATCAT	CGTGGTGATGATGGTGATGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG		

RESULT 2

US-10-205-823-412
; Sequence 412, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-412

Query Match 97.5%; Score 1034.4; DB 15; Length 1141;
Best Local Similarity 99.2%; Pred. No. 8.6e-274;
Matches 1050; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY	4	TCCTTGGGTT	CGGGTGAAG	CGCTTGGGGGTT	CAGTGGGCCAT	GATCCCCGAGCT	GCTGG	63			
Db	1	TCCTTGGGTT	CGGGTGAAG	CGCTTGGGGGTT	C--GTGGCCAT	GATCCCCGAGCT	GCTGG	58			
QY	64	AGAACTGAAG	CGGACGGTCT	CTCTGCGAA	AACAGGCAAT	GCGGAGCTGG	AGTTTGTTC	123			
Db	59	AGAACTGAAG	CGGACAGTCT	CTCTGCGAA	AACAGGCAAT	GCGGAGCTGG	AGTTTGTTC	118			
QY	124	GATCATCAT	CGTGGTGAT	GATGGTGAT	GTTGGTGAT	CACGTGCCT	GCTGAG	183			
Db	119	GATCATCAT	CGTGGTGAT	GATGGTGAT	GTTGGTGAT	CACGTGCCT	GCTGAG	178			
QY	184	CCACTACAAG	CTGTCTGC	ACGGTCTT	CATCAGCCGG	CACAGCCAGG	GGCGGAGAGAGA	243			
Db	179	CCACTACAAG	CTGTCTGC	ACGGTCTT	CATCAGCCGG	CACAGCCAGG	GGCGGAGAGAGA	238			
QY	244	AGATGCCCT	GTCTCAGA	AGGATGCCT	TGTGGCCCT	CGGAGAGCA	CAGTGT	CAGGCAACGG	303		
Db	239	AGATGCCCT	GTCTCAGA	AGGATGCCT	TGTGGCCCT	CGGAGAGCA	CAGTGT	CAGGCAACGG	298		
QY	304	AATCCAGAG	CCGCGAGG	CTACGCCCG	CCCTCGGCC	CCACCGAC	CGCCTGG	CCCGTGC	CGCC	363	
Db	299	AATCCAGAG	CCGCGAGG	CTACGCCCG	CCCTCGGCC	CCACCGAC	CGCCTGG	CCCGTGC	CGCC	358	
QY	364	CTTCGCC	CCAGCGG	AGCGCTT	CCAACGCTT	CCAGCCCC	ACCTAT	CCGTTAC	CTGCAG	CACGA	423
Db	359	CTTCGCC	CCAGCGG	AGCGCTT	CCAACGCTT	CCAGCCCC	ACCTAT	CCGTTAC	CTGCAG	CACGA	418

QY 541 GGTGCGGCAACCCCAACAGAAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGCT 600
Db |||||
Db 549 GGTGCGGCAACCCCAACAGAAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGCT 608
QY 601 GGGCGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCACGTGCTACGGCAGCGG 660
Db |||||
Db 609 GGGCGGCCCTGCCCCCCCCAGCAGTAACCTCGGGCATCAGCGCACGTGCTACGGCAGCGG 668
QY 661 CGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTCTATCGGCCACTACCGGGGTC 720
Db |||||
Db 669 CGGGCGCATGGAGGGGCGCGCCACCTACAGCGAGGTCTATCGGCCACTACCGGGGTC 728
QY 721 CTCCTTCAGCACAGCAGAGCAGTGGGCGCCCTCTTGTGAGGGGACCCGGCTCCA 780
Db |||||
Db 729 CTCCTTCAGCACAGCAGAGCAGTGGGCGCCCTCTTGTGAGGGGACCCGGCTCCA 788
QY 781 CCACACACATCGCGCCCTTAGAGGCGCAGCCATCTGGAGCAAGAGAGATAAACA 840
Db |||||
Db 789 CCACACACATCGCGCCCTTAGAGGCGCAGCCATCTGGAGCAAGAGAGATAAACA 848
QY 841 GAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGCTGCGTAGGTGAAAAG 900
Db |||||
Db 849 GAAAGGACACCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGCTGCGTAGGTGAAAAG 908
QY 901 GCAGAACACTCCGCGCTTCTTAGAGAGGAGTGAGAGGAAGGGGGGGCGCAGCAACGC 960
Db |||||
Db 909 GCAGAACACTCCGCGCTTCTTAGAGAGGAGTGAGAGGAAGGGGGGGCGCAGCAACGC 968
QY 961 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTC 1020
Db |||||
Db 969 ATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTC 1028
QY 1021 TGAATGCAAGCTAAGAGAGCTTGCAAAAAAAAAAAAAA 1058
Db |||||
Db 1029 TGAATGCAAGCTAAGAGAGCTTGCAAAAAAAAAAAAAA 1066

RESULT 7
US-09-796-753-55
; Sequence 55, Application US/09796753
; Publication No. US2003002798A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30

; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 55
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)...(761)
US-09-796-753-55

Query Match 90.7%; Score 961.8; DB 11; Length 969;
Best Local Similarity 99.8%; Pred. No. 6.9e-254;
Matches 963; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 97 GGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTGCTGCTGATGATGATGATGAT 156
Db |||||
Db 2 GGAGATGGCGGAGCTGGAGTTTGTTCAGATCATCATCTGCTGCTGATGATGATGATGAT 61
QY 157 GGTGGTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTCTTCATCAG 216
Db |||||
Db 62 GGTGGTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTCTTCATCAG 121
QY 217 CCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCTCTGGGCC 276
Db |||||
Db 122 CCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCAGAAAGGATGCTCTGGGCC 181
QY 277 CTCGGAGAGCACAGTGTCAAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCG 336
Db |||||
Db 182 CTCGGAGAGCACAGTGTCAAGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCTCG 241
QY 337 GCCCACCAGCCGCTGCGCTGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCA 396
Db |||||
Db 242 GCCCACCAGCCGCTGCGCTGCGCCCTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCA 301
QY 397 GCCCACCATATCCGTACTCTGCAGCACGAGATCGACTGCGGCCCCACCATCTCGTGTGAGA 456
Db |||||
Db 302 GCCCACCATATCCGTACTCTGCAGCACGAGATCGACTGCGGCCCCACCATCTCGTGTGAGA 361
QY 457 CGGGGAGGAGCCCCACCTACAGGGCCCCCTGACCCCTCCAGCTTCGGGACCCCGAGCA 516
Db |||||
Db 362 CGGGGAGGAGCCCCACCTACAGGGCCCCCTGACCCCTCCAGCTTCGGGACCCCGAGCA 421
QY 517 GCAGCTGGAACCTGAACCGGGAGTGGTGCAGCGACCCCGGAGGAGGAGGAGGAGGAGGAG 576
Db |||||
Db 422 GCAGCTGGAACCTGAACCGGGAGTGGTGCAGCGACCCCGGAGGAGGAGGAGGAGGAGGAG 481
QY 577 TGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTTCGCCCCCGGAGGAGGAGGAGGAGGAG 636

Db 482 TGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTGCCCGCCAGCAGTAACCTCGGGCAT 541
Qy 637 CAGCGCCACGTGTACTACGGCAGCGGGCGGCGCATGGAGGGCGCGCCACCTACAGCGA 696
Db 542 CAGCGCCACGTGTACTACGGCAGCGGGCGGCGCATGGAGGGCGCGCCACCTACAGCGA 601
Qy 697 GGTATCGGCGCCACTACCCGGGCTCCTCCAGCACACACATCGCGCCCTAGAGAGCGCAGCAT 756
Db 602 GGTATCGGCGCCACTACCCGGGCTCCTCCAGCACACACATCGCGCCCTAGAGAGCGCAGCAT 661
Qy 757 CTGCTGGAGGGGACCCGGCTCCACCCACACACATCGCGCCCTAGAGAGCGCAGCAT 816
Db 662 CTGCTGGAGGGGACCCGGCTCCACCCACACACATCGCGCCCTAGAGAGCGCAGCAT 721
Qy 817 CTGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGG 876
Db 722 CTGAGCAAGAGAGGATTAACAGAAAGGACACCCCTCTTAGGGTCCCCAGGGGGCGG 781
Qy 877 GGCTGGGGCTGCGTAGGTGAAGGCGAGAAACACTCCGGCGTCTTAGAAGAGGAGTGAGA 936
Db 782 GGCTGGGGCTGCGTAGGTGAAGGCGAGAAACACTCCGGCGTCTTAGAAGAGGAGTGAGA 841
Qy 937 GGAAGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATA 996
Db 842 GGAAGCGGGGGCGCAGCAACGATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATA 901
Qy 997 AATATTTACATGTGATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAA 1056
Db 902 AATATTTACATGTGATGTCTGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAA 961
Qy 1057 AAAAA 1061
Db 962 AAAAA 966

RESULT 8

US-10-241-220-119
; Sequence 119, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Prantz,Gretchen
; APPLICANT: Hillan,Kenneth J.
; APPLICANT: Phillips,Heidi
; APPLICANT: Polakis,Paul
; APPLICANT: Spencer,Susan
; APPLICANT: Williams,P.Mickey
; APPLICANT: Wu,Thomas
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 119
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-241-220-119

Query Match 90.1%; Score 956; DB 13; Length 4839;
Best Local Similarity 98.5%; Pred. No. 4e-252;
Matches 965; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 82 TCTCCTGGAAACCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 141
Db 407 TTTGTTCCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 466
Qy 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 201
Db 467 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 526

Qy 202 ACGGTCTTCATCAGCGGCGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCAGA 261
Db 527 ACGGTCTTCATCAGCGGCGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCAGA 586
Qy 262 AGGATGCCTGTGGCCTCGGAGAGCACAGTGTACGCAACGGAATCCCAGAGCCGAGGT 321
Db 587 AGGATGCCTGTGGCCTCGGAGAGCACAGTGTACGCAACGGAATCCCAGAGCCGAGGT 646
Qy 322 CTACGCCCCGCTCGGCGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 381
Db 647 CTACGCCCCGCTCGGCGCCACCGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 706
Qy 382 CTTCCACCGCTTCAGCGCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCAC 441
Db 707 CTTCCACCGCTTCAGCGCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCAC 766
Qy 442 CATCTCGCTGTACAGCGGGAGGAGCCCCCACCCTACAGGGCCCCCTGCACCTCCAGCT 501
Db 767 CATCTCGCTGTACAGCGGGAGGAGCCCCCACCCTACAGGGCCCCCTGCACCTCCAGCT 826
Qy 502 TCGGACCCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCACCCCCCAACAG 561
Db 827 TCGGACCCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTGGTGGCGGCACCCCCCAACAG 886
Qy 562 AACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGCGGCCCTGCCCCCCCAG 621
Db 887 AACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGCGGCCCTGCCCCCCCAG 946
Qy 622 CAGTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGGGCGGCGCATGGAGGGCGGCC 681
Db 947 CAGTAACTCGGGCATCAGCGCCACCTGCTACGGCAGCGGGCGGCGCATGGAGGGCGGCC 1006
Qy 682 GCCCACCTACAGCGAGTTCATCGGCCACTACCCGGGGTCTCTTCCAGCACCAGAGAG 741
Db 1007 GCCCACCTACAGCGAGTTCATCGGCCACTACCCGGGGTCTCTTCCAGCACCAGAGAG 1066
Qy 742 CAGTGGGGCCCTCTCTGCTGGAGGGGACCCGGTCCACCAACACACACATCGCGCCCT 801
Db 1067 CAGTGGGGCCCTCTCTGCTGGAGGGGACCCGGTCCACCAACACACATCGCGCCCT 1126
Qy 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTTAGGG 861
Db 1127 AGAGAGCGCAGCCATCTGGAGCAAGAGAGAGATTAACAGAAAGGACACCTCTTAGGG 1186
Qy 862 TCCCAGGGGGCGGCTGGGCTGCGTAGGTGAAAGGAGAGAGAGAGAGAGAGAGAGAG 921
Db 1187 TCCCAGGGGGCGGCTGGGCTGCGTAGGTGAAAGGAGAGAGAGAGAGAGAGAGAGAG 1246
Qy 922 AGAAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
Db 1247 AGAAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1306
Qy 982 ACCTCCCTGTGTATAAATATTACATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1041
Db 1307 ACCTCCCTGTGTATAAATATTACATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1366
Qy 1042 CTTGCAAAAAA 1061
Db 1367 CTTGCAAAAAA 1386

RESULT 9

US-10-269-909-84
; Sequence 84, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909

```

; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-84

```

Query Match		90.1%;	Score 956;	DB 13;	Length 4839;
Best Local Similarity		98.5%;	Pred. No. 4e-252;		
Matches 965;		Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;
QY	82	TCTCCTGCGAAAC	CAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	141	
DB	407	TTTGTTCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	466		
QY	142	GGTGATGATGGTGTGGTGGTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGC	201		
DB	467	GGTGATGATGGTGTGGTGGTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGC	526		
QY	202	ACGGTCCTTCATCAGCCGGCACAGCCAGGGGGGAGGAGAGAAAGATGCCCTGTCTCAG	261		
DB	527	ACGGTCCTTCATCAGCCGGCACAGCCAGGGGGGAGGAGAGAAAGATGCCCTGTCTCAG	586		
QY	262	AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGAGGT	321		
DB	587	AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGAGGT	646		
QY	322	CTACGCCCCGCTCGGCCCAACGACCGCCTGGCGGTGCGGCCCTTCGCCCAAGCGGAGCG	381		
DB	647	CTACGCCCCGCTCGGCCCAACGACCGCCTGGCGGTGCGGCCCTTCGCCCAAGCGGAGCG	706		
QY	382	CTTCCACCGCTTCAGACCCACCTATCCGTACCTGCAGCACAGATCGACCTGCCGCCAC	441		
DB	707	CTTCCACCGCTTCAGACCCACCTATCCGTACCTGCAGCACAGATCGACCTGCCGCCAC	766		
QY	442	CATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCAGGGCCCCCTGCACCCCTCCAGCT	501		
DB	767	CATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCAGGGCCCCCTGCACCCCTCCAGCT	826		
QY	502	TCGGGACCCCGAGCAGCAGCTGGAACCTGAAACGGGAGTCGGTCGCGCAACCCCAACAG	561		
DB	827	TCGGGACCCCGAGCAGCAGCTGGAACCTGAAACGGGAGTCGGTCGCGCAACCCCAACAG	886		
QY	562	AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGTGGCGGCCCTGCCCCGCCAG	621		
DB	887	AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGTGGCGGCCCTGCCCCGCCAG	946		
QY	622	CAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCC	681		
DB	947	CAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCC	1006		
QY	682	GCCCAACCTACAGCGAGGTTCATCGGCCACTACCCGGGTCTCTTCCAGCACCCAGCAGAG	741		
DB	1007	GCCCAACCTACAGCGAGGTTCATCGGCCACTACCCGGGTCTCTTCCAGCACCCAGCAGAG	1066		
QY	742	CAGTGGGCCGCCCTCTTGTGCTGGAGGGGACCCGGCTCCACCAACACATCGCGGCCCT	801		
DB	1067	CAGTGGGCCGCCCTCTTGTGCTGGAGGGGACCCGGCTCCACCAACACATCGCGGCCCT	1126		
QY	802	AGAGAGCGCAGCCATCTCTGGAGCAAAGAGAGGATAAACAGAAAGGACACCTCTTAGGG	861		
DB	1127	AGAGAGCGCAGCCATCTGGAGCAAAGAGAGGATAAACAGAAAGGACACCTCTTAGGG	1186		
QY	862	TCCCCAGGGGGCCGGGCTGGGGCTGCGTAGGTGAAAGGCAGAAACACTCCGGCGTTCTT	921		
DB	1187	TCCCCAGGGGGCCGGGCTGGGGCTGCGTAGGTGAAAGGCAGAAACACTCCGGCGTTCTT	1246		

QY	922	AGAAGAGGAGTGAGAGGAAGGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCC	981
Db	1247	AGAAGAGGAGTGAGAGGAAGGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCC	1306
QY	982	ACCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAG	1041
Db	1307	ACCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGAG	1366
QY	1042	CTTGCAAAAAAAAAAAAAA	1061
Db	1367	CTTGCAAAAAAAAAAAAAA	1386

RESULT 10
US-10-269-909-85
; Sequence 85, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIREBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-85

	Query Match	90.1%;	Score 956;	DB 13;	Length 4839;
	Best Local Similarity	98.5%;	Pred. No. 4e-252;		
	Matches 965;	Conservative	0;	Mismatches 15;	Indels 0;
				Gaps	0;
QY	82	TCTCTGCGAAAC	CAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	141	
DB	407	TTTGTTCAGAGCATGGAGATCACGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGT	466		
QY	142	GGTGATGATGTTGTTGTTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGC	201		
DB	467	GGTGATGATGTTGTTGTTGATCACGTGCCTGCTGAGCCACTACAAGCTGTCTGC	526		
QY	202	ACGGTCCTTCATCAGCCGGCACAGCCAGGGGCGGAGAGAGAAGATGCCCTGTCTCTCAGA	261		
DB	527	ACGGTCCTTCATCAGCCGGCACAGCCAGGGGCGGAGAGAGAAGATGCCCTGTCTCTCAGA	586		
QY	262	AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCACAGAGCCGCAGGT	321		
DB	587	AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAAATCCACAGAGCCGCAGGT	646		
QY	322	CTACGCCCGCCTCGGCCCAACCGACCGCCTGGCCGTGCCGCCCTTCGCCACGCGGAGCG	381		
DB	647	CTACGCCCGCCTCGGCCCAACCGACCGCCTGGCCGTGCCGCCCTTCGCCACGCGGAGCG	706		
QY	382	CTTCCACCGTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCAC	441		
DB	707	CTTCCACCGTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCAC	766		
QY	442	CATCTCGCTGTTCAGACGGGGAGGAGCCCACTACCGAGGCCCTTGACACCTCTCCAGCT	501		
DB	767	CATCTCGCTGTTCAGACGGGGAGGAGCCCACTACCGAGGCCCTTGACACCTCTCCAGCT	826		
QY	502	TCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCCGGTCCGCGCACCCCCCAACAG	561		
DB	827	TCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCCGGTCCGCGCACCCCCCAACAG	886		

QY 562 AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTGCCCCCAG 621
DB 887 AACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGGCGGCCCTGCCCCCAG 946
QY 622 CAGTAACTCGGGCATCAGCGCACAGTGTCTACGGCAGCGGGCGCATGGAGGGCGGCC 681
DB 947 CAGTAACTCGGGCATCAGCGCACAGTGTCTACGGCAGCGGGCGCATGGAGGGCGGCC 1006
QY 682 GCCACCTACAGCGAGGTATCGGCCACTACCGGGGTCTCTCTCCAGCACCGAGAG 741
DB 1007 GCCACCTACAGCGAGGTATCGGCCACTACCGGGGTCTCTCTCCAGCACCGAGAG 1066
QY 742 CAGTGGCGGCCCTCTCTGCTGGAGGGGACCCCGCTCCACACACACATCGCGGCCCT 801
DB 1067 CAGTGGCGGCCCTCTCTGCTGGAGGGGACCCCGCTCCACACACACATCGCGGCCCT 1126
QY 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGG 861
DB 1127 AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGGG 1186
QY 862 TCCCGAGGGGGCCGGCTGGGGCTGCGTAGGTGAAAGGAGAACTCCCGCTCTT 921
DB 1187 TCCCGAGGGGGCCGGCTGGGGCTGCGTAGGTGAAAGGAGAACTCCCGCTCTT 1246
QY 922 AGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCGATCGTGTGGCCCTCC 981
DB 1247 AGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCGATCGTGTGGCCCTCC 1306
QY 982 ACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAG 1041
DB 1307 ACCTCCCTGTGTATAAATATTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAG 1366
QY 1042 CTGCAAAAAA 1061
DB 1367 CTGCAAAAAA 1386

RESULT 11

US-09-934-249-1
; Sequence 1, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landeschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (413)...(1273)
US-09-934-249-1

Query Match 75.5%; Score 800.6; DB 10; Length 1321;
Best Local Similarity 98.3%; Pred. No. 1.3e-209;
Matches 809; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 82 TCTCCTGCGAAACAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 141
DB 499 TTTGTTCCAGAGCATGGAGATCAGGAGCTGGAGTTTGTTCAGATCATCATCGTGGT 558

QY 142 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 201
DB 559 GGTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 618
QY 202 ACGGTCTTTCATCAGCGCGCACAGCCAGCGGGCGGAGAGAGAGAGATGCCCTGTCTCAGA 261
DB 619 ACGGTCTTTCATCAGCGCGCACAGCCAGCGGGCGGAGAGAGAGATGCCCTGTCTCAGA 678
QY 262 AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCTAGCAACCGGAATCCCAGAGCCGAGGT 321
DB 679 AGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCTAGCAACCGGAATCCCAGAGCCGAGGT 738
QY 322 CTAGCCCCCGCTCGGCCACCGACCGCTGGCGTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 381
DB 739 CTAGCCCCCGCTCGGCCACCGACCGCTGGCGTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 798
QY 382 CTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGACGACGAGATCGACCTGCCGCCAC 441
DB 799 CTTCCACCGCTTCCAGCCCCACCTATCCGTACCTGACGACGAGATCGACCTGCCGCCAC 858
QY 442 CATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCGAGGGCCCTGACCCCTCCAGCT 501
DB 859 CATCTCGCTGTACAGCGGGAGGAGCCCCACCTACCGAGGGCCCTGACCCCTCCAGCT 918
QY 502 TCGGACCCCCGAGCAGAGCTGGAACCTGAACCGGGAGTGGTGGCGCACCCCCCAACAG 561
DB 919 TCGGACCCCCGAGCAGAGCTGGAACCTGAACCGGGAGTGGTGGCGCACCCCCCAACAG 978
QY 562 AACCATCTTCGACAGTGTGATGGATAGTGCAGGGTGGGGGGGGGGGGGGGGGGGGGGGG 621
DB 979 AACCATCTTCGACAGTGTGATGGATAGTGCAGGGTGGGGGGGGGGGGGGGGGGGGGGGG 1038
QY 622 CAGTAACTCGGGCATCAGCGCCACAGTGTCTACGGCAGCGGGCGCATGGAGGGCGGCC 681
DB 1039 CAGTAACTCGGGCATCAGCGCCACAGTGTCTACGGCAGCGGGCGCATGGAGGGCGGCC 1098
QY 682 GCCACCTACAGCGAGGTCTCGGGCAGTCCCGGGGACCCGGTCCACACACACATCGCGCCCT 741
DB 1099 GCCACCTACAGCGAGGTCTCGGGCAGTCCCGGGGACCCGGGTCTCTCTCCAGCACCGAG 1158
QY 742 CAGTGGCGGCCCTCTCTGCTGGAGGGGACCCGGTCCACACACACATCGCGCCCT 801
DB 1159 CAGTGGCGGCCCTCTCTGCTGGAGGGGACCCGGTCCACACACACATCGCGCCCT 1218
QY 802 AGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTTAGGG 861
DB 1219 AGAGAGCGCAGCCATCTGGAGCAAGAGAGAGGATAAACAGAAAGGACACCTCTTAGGG 1278
QY 862 TCCCCAGGGGGCGGGCTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGG 904
DB 1279 TCCCCAGGGGGCGGGCTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGG 1321

RESULT 12

US-10-241-220-45
; Sequence 45, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120


```
; SEQ ID NO 45
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-241-220-45

Query Match      74.9%; Score 795.2; DB 13; Length 806;
Best Local Similarity 99.6%; Pred. No. 3.5e-208;
Matches 797; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 75 CGGACGGTCTCTTGCAGAACCCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATATCA 134
Db 7 CGGACAGTCTCTTGCAGAACCCAGGCAATGGCGGAGCTGGAGTTTGTTCAGATCATATCA 66

QY 135 TCGTGGTGGTATGATGGTATGGTGGTGGTATCACGTGCTGCTGAGCCACTACAAGC 194
Db 67 TCGTGGTGGTATGATGGTATGGTGGTGGTATCACGTGCTGCTGAGCCACTACAAGC 126

QY 195 TGTCTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGT 254
Db 127 TGTCTGCACGGTCTCTTCATCAGCCGGCACAGCCAGGGCGGAGGAGAGAAGATGCCCTGT 186

QY 255 CCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGC 314
Db 187 CCTCAGAAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGC 246

QY 315 CGCAGGTCTACGCCCCCGCTCGGCCACCGACCGCTGGCCGTGGCCGCTTCCGCCCAGC 374
Db 247 CGCAGGTCTACGCCCCCGCTCGGCCACCGACCGCTGGCCGTGGCCGCTTCCGCCCAGC 306

QY 375 GGGAGCGGTTCCACCGTTCCAGCCCACTATCCGTACTTGAGAGCACGAGATCGACCTGC 434
Db 307 GGGAGCGGTTCCACCGTTCCAGCCCACTATCCGTACTTGAGAGCACGAGATCGACCTGC 366

QY 435 CGCCCAACCATCTCGCTGTAGACGGGGAGGAGCCCCACCTACAGGGCCCTTGCAACC 494
Db 367 CGCCCAACCATCTCGCTGTAGACGGGGAGGAGCCCCACCTACAGGGCCCTTGCAACC 426

QY 495 TCCAGCTTCGGGACCCCGAGCAGAGCTGGAACTGAAACCGGAGTCGGTGC CGCGCACCCC 554
Db 427 TCCAGCTTCGGGACCCCGAGCAGAGCTGGAACTGAAACCGGAGTCGGTGC CGCGCACCCC 486

QY 555 CAAACAGAACCATCTTCGACAGTGAACCTGATGGATAGTGCAGAGTGGCGGCCCTGCC 614
Db 487 CAAACAGAACCATCTTCGACAGTGAACCTGATGGATAGTGCAGAGTGGCGGCCCTGCC 546

QY 615 CCCCAGCAGTAACCTCGGGCATCAGGCCACAGTGTACGGCAGCGGGCGGCATGGAGG 674
Db 547 CCCCAGCAGTAACCTCGGGCATCAGGCCACAGTGTACGGCAGCGGGCGGCATGGAGG 606

QY 675 GGCCGCGCCCACTACAGCGAGGTATCGGCCACTACCGGGGTCTCTTCCAGCACC 734
Db 607 GGCCGCGCCCACTACAGCGAGGTATCGGCCACTACCGGGGTCTCTTCCAGCACC 666

QY 735 AGCAGAGCAGTGGGCGCGCCCTCTTGTGGAGGGGACCCGGTCCACCAACACACATCG 794
Db 667 AGCAGAGCAGTGGGCGCGCCCTCTTGTGGAGGGGACCCGGTCCACCAACACACATCG 726

QY 795 CGCCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCCCTC 854
Db 727 CGCCCCTAGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCCCTC 786

QY 855 TCTAGGTCCCCAGGGGGC 874-
Db 787 TCTAGGTCCCCAGAGGGC 806
```

RESULT 13
US-10-390-045-2
; Sequence 2, Application US/10390045
; Publication No. US2003017013A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV

```
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-390-045-2
```

```
Query Match      71.4%; Score 757.4; DB 13; Length 759;
Best Local Similarity 99.9%; Pred. No. 8.1e-198;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 101 ATGGCGGAGCTGGAGTTGTTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 160
Db 1 ATGGCGGAGCTGGAGTTGTTTCAGATCATCATCATCATCATCATCATCATCATCATCAT 60

QY 161 GTGGTGATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGG 220
Db 61 GTGGTGATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTTCATCAGCCGG 120

QY 221 CACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCCTCAGAAAGGATGCCCTGTGGCCCTCG 280
Db 121 CACAGCCAGGGCGGAGGAGAGAAGATGCCCTGTCTCCTCAGAAAGGATGCCCTGTGGCCCTCG 180

QY 281 GAGAGCACAGTGTCTAGGCAACGGAATCCCAGAGCCCGCAGGTCCTACGCCCCCGCTCGGCC 340
Db 181 GAGAGCACAGTGTCTAGGCAACGGAATCCCAGAGCCCGCAGGTCCTACGCCCCCGCTCGGCC 240

QY 341 ACCGACCGCCTGGCCGTGCGCCCTTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCAGAGCCC 400
Db 241 ACCGACCGCCTGGCCGTGCGCCCTTTCGCCCCAGCGGGAGCGCTTCCACCGCTTCAGAGCCC 300

QY 401 ACCTATCCGTACTCTGAGCACGAGATCGACCTGCGGCCCAACCATCTCGCTGTACAGACGGG 460
Db 301 ACCTATCCGTACTCTGAGCACGAGATCGACCTGCCACCCACCATCTCGCTGTACAGACGGG 360

QY 461 GAGGAGCCCCCACCCTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAG 520
Db 361 GAGGAGCCCCCACCCTACAGGGCCCCCTGCACCTCCAGCTTCGGGACCCCGAGCAGCAG 420

QY 521 CTGGAACTGAACCGGGAGTCGGTGGCGGCACCCCCCAAAACAGAACCATCTTCGACACTGAC 580
Db 421 CTGGAACTGAACCGGGAGTCGGTGGCGGCACCCCCCAAAACAGAACCATCTTCGACACTGAC 480

QY 581 CTGATGGATAGTGCACGGCTGGGCGGCCCTTGCCTCCCCCAGCAGTAACTCGGGCATCAGC 640
Db 481 CTGATGGATAGTGCACGGCTGGGCGGCCCTTGCCTCCCCCAGCAGTAACTCGGGCATCAGC 540

QY 641 GCCACGTGCTACGGCAGCGCGGGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTC 700
Db 541 GCCACGTGCTACGGCAGCGCGGGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTC 600

QY 701 ATCGGCCACTACCCGGGTCTCTCTTCCAGCACACAGCAGAGCAGTGGGCCCGCTCTCTTG 760
Db 601 ATCGGCCACTACCCGGGTCTCTCTTCCAGCACACAGCAGAGCAGTGGGCCCGCTCTCTTG 660

QY 761 CTGGAGGGGACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGG 820
Db 661 CTGGAGGGGACCCGGCTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGG 720
```

QY	622	CAGTAACTCGGGCATCAGCCCACTGTCTACGGCAGCGGGCGCATGGAGGGCGCCG	681
Db	627	CAGTAACTCGGGCATCAGCGCACTGTCTACCGACCGGGCGGGCGCATGGAGGGCGCCG	686

QY	682	GCCACCTACAGCGAGGTCATCGGCCACTACCCGGGGTCTCTTCCAGCACCGACGAG	746
Db	687	GCCACCTACAGCGAGGTCATCGGCCACTACCCGGGGTCTCTTCCAGCACCGACGAG	746
QY	742	CAGTGGGGCCGCCCTCTTGCTGGAGGGGACCCGGCTCCACACACACATCGCGCCCT	801
Db	747	CAGTGGGGCCGCCCTCTTGCTGGAGGGGACCCGGCTCCACACACACATCGCGCCCT	806
QY	802	AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCCCTCTC	856
Db	807	AGAGAGCGCAGCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCCCTCTC	861

RESULT 15
US-10-000-256A-32
; Sequence 32, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION: .

```

; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01

```

Query Match	70.9%;	Score 752.6;	DB 10;	Length 861;
Best Local Similarity	98.2%;	Pred. No. 1.7e-196;		
Matches 761;	Conservative	0;	Mismatches 14;	Indels 0;
				Gaps 0;

```

; TYPE: DNA
; ORGANISM: Homo sapien
US-10-000-256A-32

Query Match      68.7%;      Score 728.8;      DB 15;      Length 1583;
Best Local Similarity 98.9%;      Pred. NO. 6.8e-190;
Matches 744; Conservative 0; Mismatches 7; Indels 1; Gaps 1

```

QY	311	GAGCGCGAGGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGAC	430
Db	825	GCGCCGCGAGGCTTACGCCCGCCCGCTCGGCCCAACCGACCGCCTGGCGTGCCGCCCTTCGCC	884
QY	371	CAGCGGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGAC	430
Db	885	CAGCGGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGAC	944

[illegible]

Accession	Sequence	Position
Db	1065 CCCCCAACAGAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGCGGGCCCC	112
Qy	611 TGCCCCCCAGCAGTAACCTCGGGCATCAGGCCACGTGCTACGGCAGCGGGCGGCATG	670
Db	1125 TGCCCCCCCCAGCAGTAACCTCGGGCATCAGGCCACGTGCTACGGCAGCGGGCGGCATG	118

D_b 1185 GAGGGGCGCGCCACCTACAGCGAGGTCATCGGCCACTACCGGGGTCTCTCTCCAG 124

Qy	731	CACCAGCAGAGCAGTGGGCCGCCCTCCTTGCTGGAGGGACCCGGCTCCACCACACACAC	790
Db	1245		
Qy	791	ATCGCGCCCTAGAGAGCGCAGGCATCTGGAGCAAGAGAGGATAAACAGAAAGGACAC	850
Db	1305		
Qy	851	CCTCTCTAGGGTCCCCAGGGGGCCGGGCTGGGGCTGCCGTAGGTGAAAAGGCAGAACACT	910
Db	1365		
Qy	911	CCGCGCTTCTTAGAAGAGGAGTGAGAGGAAGCGGGGGCGCGCAGCAACGCAT-CGTGTGG	969
Db	1425		
Qy	970	CCCTCCCTCCACCTCCCTGTGTATAAAATATTTACATGTGATGTCTGGTCTGAATGCAC	1029
Db	1485		
Qy	1030	AAGCTAAGAGAGCTTGCAAAAAAAAAAAAAAAAAA	1061
Db	1545		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM protein - protein search, using sw model
Run on: December 4, 2003, 19:50:19 ; Search time 31 Seconds
(without alignments)
1511.867 Million cell updates/sec

Title: US-09-857-826B-17
Perfect score: 1348
Sequence: 1 MAELFVQIIIVVMVMV.....PLESAAIWSKEKDKQKGHPL 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1348	100.0	252	11 US-09-796-753-56	Sequence 56, Appl
2	1348	100.0	252	12 US-10-241-220-100	Sequence 100, App
3	1348	100.0	252	12 US-10-241-220-101	Sequence 101, App
4	1348	100.0	252	12 US-10-301-822-209	Sequence 209, App
5	1348	100.0	252	12 US-09-821-812-3	Sequence 3, Appli
6	1348	100.0	252	12 US-10-390-045-3	Sequence 3, Appli
7	1348	100.0	252	15 US-10-205-823-413	Sequence 413, App
8	1340	99.4	287	10 US-09-934-249-2	Sequence 2, Appli
9	1340	99.4	287	12 US-10-241-220-120	Sequence 120, App
10	1332	98.8	249	12 US-10-390-045-11	Sequence 11, Appl
11	1172.5	87.0	274	10 US-09-934-249-13	Sequence 13, Appl
12	1066.5	79.1	217	11 US-09-796-753-58	Sequence 58, Appl
13	1065.5	79.0	241	15 US-10-000-256A-158	Sequence 158, App
14	845.5	62.7	244	12 US-10-390-045-12	Sequence 12, Appl
15	845.5	62.7	269	12 US-10-094-749-1836	Sequence 1836, Ap

16	842.5	62.5	306	12	US-10-251-598-3	Sequence 0, Appli
17	791.5	58.7	288	12	US-10-251-598-4	Sequence 0, Appli
18	104	7.7	830	10	US-09-870-759-134	Sequence 134, App
19	104	7.7	830	11	US-09-842-758-57	Sequence 57, Appl
20	104	7.7	830	12	US-09-751-708A-134	Sequence 134, App
21	103	7.6	1318	12	US-10-200-562-197	Sequence 197, App
22	103	7.6	1318	12	US-10-237-551-197	Sequence 197, App
23	100	7.4	830	10	US-09-870-759-140	Sequence 140, App
24	100	7.4	830	12	US-09-751-708A-140	Sequence 140, App
25	100	7.4	2440	12	US-10-341-434-236	Sequence 236, App
26	99.5	7.4	721	14	US-10-086-464-5	Sequence 5, Appli
27	99	7.3	312	10	US-09-789-054A-22	Sequence 22, Appl
28	99	7.3	427	15	US-10-050-704-273	Sequence 273, App
29	99	7.3	483	15	US-10-050-704-272	Sequence 272, App
30	99	7.3	671	11	US-09-983-000A-14	Sequence 14, Appl
31	98	7.3	263	10	US-09-974-298-110	Sequence 110, App
32	98	7.3	2429	12	US-10-288-798-8	Sequence 8, Appli
33	97	7.2	527	10	US-09-712-363-156	Sequence 156, App
34	97	7.2	527	16	US-10-080-170-348	Sequence 348, App
35	96.5	7.2	366	12	US-10-192-381-20	Sequence 20, Appl
36	95.5	7.1	707	11	US-09-919-039-278	Sequence 278, App
37	95.5	7.1	859	15	US-10-153-668-324	Sequence 324, App
38	95	7.0	649	11	US-09-759-130B-333	Sequence 333, App
39	95	7.0	649	12	US-10-188-495-63	Sequence 63, Appl
40	95	7.0	649	15	US-10-189-123-63	Sequence 63, Appl
41	95	7.0	671	11	US-09-759-130B-331	Sequence 331, App
42	95	7.0	671	12	US-10-188-495-61	Sequence 61, Appl
43	95	7.0	671	15	US-10-189-123-61	Sequence 61, Appl
44	95	7.0	876	12	US-10-240-154-14	Sequence 14, Appl
45	95	7.0	876	15	US-10-060-036-4553	Sequence 4553, Ap

ALIGNMENTS

RESULT 1
US-09-796-753-56
; Sequence 56, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23

684280

; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 56
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-753-56

Query Match 100.0%; Score 1348; DB 11; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
|
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60

Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
|
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
|
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180

Qy 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
|
Db 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

Qy 241 SKEKDKQKGHPL 252
|
Db 241 SKEKDKQKGHPL 252

RESULT 2
US-10-241-220-100
; Sequence 100, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220

; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 100
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-100

Query Match 100.0%; Score 1348; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
|
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60

Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
|
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
|
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180

Qy 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
|
Db 181 ATCYGSGRMEGPPPTYSEVIGHYPCSSFHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240

Qy 241 SKEKDKQKGHPL 252
|
Db 241 SKEKDKQKGHPL 252

RESULT 3
US-10-241-220-101
; Sequence 101, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5010R1-US
; CURRENT APPLICATION NUMBER: US/10/241,220
; CURRENT FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 101
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-241-220-101

Query Match 100.0%; Score 1348; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
|
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60

Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120
|
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDG 120

Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
|
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180

Qy 181 ATCYSGGRMEGPPPTTYSEVIGHYPCSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGRMEGPPPTTYSEVIGHYPCSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 4
US-10-301-822-209
; Sequence 209, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-020P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-209

Query Match 100.0%; Score 1348; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180
Qy 181 ATCYSGGRMEGPPPTTYSEVIGHYPCSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGRMEGPPPTTYSEVIGHYPCSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 5
US-09-821-812-3
; Sequence 3, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:

; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; TITLE OF INVENTION: Nucleic Acids
; FILE REFERENCE: R-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-812-3

Query Match 100.0%; Score 1348; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
Db 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
Db 61 ESTVSGNGIPEQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLSDG 120
Qy 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180
Db 121 EEPYPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGCGPCPPSSNSGIS 180
Qy 181 ATCYSGGRMEGPPPTTYSEVIGHYPCSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db 181 ATCYSGGRMEGPPPTTYSEVIGHYPCSSFQHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 6
US-10-390-045-3
; Sequence 3, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995 0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-390-045-3

Query Match 100.0%; Score 1348; DB 12; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRRREDALSSEGCLWPS 60

Db 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTTISLSDG 120
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTTISLSDG 120
Qy 121 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
Db 121 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
Qy 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAIWI 240
Db 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAIWI 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 7
US-10-205-823-413
; Sequence 413, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 413
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-413

Query Match 100.0%; Score 1348; DB 15; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.4e-111;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTTISLSDG 120
Db 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTTISLSDG 120
Qy 121 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180

Db 121 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
Qy 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAIWI 240
Db 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAIWI 240
Qy 241 SKEKDQKQGHPL 252
Db 241 SKEKDQKQGHPL 252

RESULT 8
US-09-934-249-2
; Sequence 2, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-934-249-2

Query Match 99.4%; Score 1340; DB 10; Length 287;
Best Local Similarity 99.2%; Pred. No. 5.6e-110;
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MALEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db 36 ITELEFVQIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 95
Qy 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTTISLSDG 120
Db 96 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLOHEIDLPTTISLSDG 155
Qy 121 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 180
Db 156 EEPYPYQGPCTQLQRLDPEQQLNRESVRAPPNRTIFDSDLMDSARLGCGPCPPSSNSGIS 215
Qy 181 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAIWI 240
Db 216 ATCYSGGGRMEGPPPTTYSEVIGHYPGSSFQHQSSGPPSLLEGTRLHHTHIAPLESAIWI 275
Qy 241 SKEKDQKQGHPL 252
Db 276 SKEKDQKQGHPL 287

RESULT 9
US-10-241-220-120
; Sequence 120, Application US/10241220
; Publication No. US20030148408A1
; GENERAL INFORMATION:
; APPLICANT: Frantz, Gretchen
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Phillips, Heidi
; APPLICANT: Polakis, Paul
; APPLICANT: Spencer, Susan
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF TUMOR
FILE REFERENCE: P5010R1-US
CURRENT APPLICATION NUMBER: US/10/241,220
CURRENT FILING DATE: 2002-12-13
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 120
LENGTH: 287
TYPE: PRT
ORGANISM: Homo sapiens
US-10-241-220-120

Query Match 99.4%; Score 1340; DB 12; Length 287;
Best Local Similarity 99.2%; Pred. No. 5.6e-110;
Matches 250; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db :
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Db :
QY 96 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 155
QY 121 EEPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 180
Db :
QY 156 EEPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 215
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db :
QY 216 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 275
QY 241 SKEKDQKQGHPL 252
Db :
QY 276 SKEKDQKQGHPL 287

RESULT 10
US-10-390-045-11
Sequence 11, Application US/10390045
Publication No. US20030170713A1
GENERAL INFORMATION:
APPLICANT: SRIVASTAVA, SHIV
APPLICANT: MOUL, JUDD W.
APPLICANT: XU, LINDA L.
APPLICANT: SEGAWA, TAKEHIKO
TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
FILE REFERENCE: 04995.0057-00000
CURRENT APPLICATION NUMBER: US/10/390,045
CURRENT FILING DATE: 2003-03-18
PRIOR APPLICATION NUMBER: US/09/769,482
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,772
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,045
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 11
LENGTH: 249
TYPE: PRT
ORGANISM: Homo sapiens
US-10-390-045-11

Query Match 98.8%; Score 1332; DB 12; Length 249;
Best Local Similarity 100.0%; Pred. No. 2.4e-109;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 AELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 61
Db :
QY 1 AELEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60

QY 62 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 121
Db :
QY 61 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Db :
QY 122 EEPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 181
Db :
QY 121 EEPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 180
Db :
QY 182 TCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 241
Db :
QY 181 TCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
QY 242 SKEKDQKQGHPL 250
Db :
QY 241 SKEKDQKQGHPL 249

RESULT 11
US-09-934-249-13
Sequence 13, Application US/09934249
Patent No. US20020115081A1
GENERAL INFORMATION:
APPLICANT: Lee, Richard T.
APPLICANT: Landschulz, Katherine T.
APPLICANT: Turi, Thomas G.
APPLICANT: Thompson, John F.
APPLICANT: Kennedy, Scott P.
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
FILE REFERENCE: P0738/7001/ERP/KA
CURRENT APPLICATION NUMBER: US/09/934,249
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 60/227,159
PRIOR FILING DATE: 2000-08-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13
LENGTH: 274
TYPE: PRT
ORGANISM: Mus Musculus
US-09-934-249-13

Query Match 87.0%; Score 1172.5; DB 10; Length 274;
Best Local Similarity 88.1%; Pred. No. 2.9e-95;
Matches 222; Conservative 11; Mismatches 10; Indels 9; Gaps 3;

QY 1 MALEFVQIIIIIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 60
Db :
QY 32 ITELEFVQIVIVVMVMVVTCLLSHYKLSARSFISRHSGRRREDALSSEGCLWPS 91
QY 61 ESTVSGNGIPEPQVYAPRPTDRLAVPPFAQRFHRFQPTYPYLQHEIDLPTISLSDG 120
Db :
QY 92 ESTVSG-GMPEPQVYAPRPTDRLAVPPFIQRS---RFQPTYPYLQHEIALPPTISLSDG 147
QY 121 EEPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLMSARLGGPCPPSSNSGIS 180
Db :
QY 148 EEPYQGPCTQLRDPQQLNRESVRAPPNRTIFDSDLIDSTMLGGPCPPSSNSGIS 207
QY 181 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 240
Db :
QY 208 ATCYGSGGRMEGPPPTYSEVIGHYPGSSFOHQSSGPPSLLEGTRLHHTHIAPLESAAIW 262
QY 241 SKEKDQKQGHPL 252
Db :
QY 263 NKEKEKQKQGHPL 274

RESULT 12
US-09-796-753-58
Sequence 58, Application US/09796753
Publication No. US20030027998A1
GENERAL INFORMATION:

APPLICANT: McCarthy, Sean A.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-227-999
CURRENT APPLICATION NUMBER: US/09/796,753
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 09/183,175
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 09/223,094
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/223,546
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/224,246
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/259,388
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 60/122,458
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 09/312,359
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/336,536
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: 09/342,687
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 09/345,464
PRIOR FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: 09/365,164
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/399,723
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 09/409,634
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 09/471,179
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 09/474,071
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/514,010
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: 09/516,745
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 09/572,002
PRIOR FILING DATE: 2000-05-14
PRIOR APPLICATION NUMBER: 09/597,993
PRIOR FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: 09/599,596
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 09/630,334
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: 09/606,565
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/606,317
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/665,666
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 09/677,751
PRIOR FILING DATE: 2000-09-30
NUMBER OF SEQ ID NOS: 162
SEQ ID NO 58
LENGTH: 217
TYPE: PRT
ORGANISM: Mouse
US-09-796-753-58

Query Match
Best Local Similarity 79.1%; Score 1066.5; DB 11; Length 217;
Matches 201; Conservative 88.9%; Pred. No. 4.8e-86;
7; Mismatches 9; Indels 9; Gaps 3;
QY 27 LSHYKLSARFISRHSGQRRRDLSEGLWPSESTVSGNGIPEPQVYAPRPTDRLAV 86
Db 1 LSHYKLSARFISRHSGQRRRDLSEGLWPSESTVSG-GMPEPQVYAPRPTDRLAV 59
QY 87 BRFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGPCTQLQLRDPEQQLELNRE 146

Db 60 PPFQIORS---RFQTPYLYLQHEIALPPTISLSDGEEPPPYQGPCTQLQLRDPEQQLELNRE 116
QY 147 SVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISATCYGSGGRMEGPPPTTYSEVIGHYPG 206
Db 117 SVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISATCYGSGGRMEGPPPTTYSEVIGHYPG 176
QY 207 SSFQHQSSGPPSLLGTRHLHHTHIAPLESAAIWSKEKDQKQGHPL 252
Db 177 SSFQHQSSGPPSLLGTRHLHHTHIAPLESAAIWSKEKDQKQGHPL 217
RESULT 13
US-10-000-256A-158
; Sequence 158, Application US/10000256A
; Publication No. US20030039983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-000-256A-158

Query Match 79.0%; Score 1065.5; DB 15; Length 241;
Best Local Similarity 82.9%; Pred. No. 6.6e-86;
Matches 199; Conservative 0; Mismatches 0; Indels 41; Gaps 1;
QY 54 EGCLWPSESTVSGNGIPE-----P 72
Db 1 EGCLWPSESTVSGNGIPECCWDPPCRRSSAPCPAGSSPALCSLHTGARTLPLFGGGRP 60
QY 73 QVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGPCTL 132
Db 61 QVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTTISLSDGEEPPPYQGPCTL 120
QY 133 QLRDPEQQLELNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISATCYGSGGRMEG 192
Db 121 QLRDPEQQLELNRESVRAPPNRTIFDSDLMD SARLGPCPPSSNSGISATCYGSGGRMEG 180
QY 193 PPPTYSEVIGHYPGSSFOHQSSGPPSLLGTRHLHHTHIAPLESAAIWSKEKDQKQGHPL 252
Db 181 PPPTYSEVIGHYPGSSFOHQSSGPPSLLGTRHLHHTHIAPLESAAIWSKEKDQKQGHPL 240

RESULT 14
US-10-390-045-12
; Sequence 12, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-00000
; CURRENT APPLICATION NUMBER: US/10/390,045
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772


```
US-10-094-749-1836

Query Match      62.7%; Score 845.5; DB 12; Length 269;
Best Local Similarity 67.6%; Pred. No. 1.8e-66;
Matches 169; Conservative 21; Mismatches 53; Indels 7; Gaps 4

QY   2 AELEFVQIIIIIIVVMVMVVITCLLSHYKLSARSFISRHSGQRREDALSSGGLWPSE 61
    ||||| ||||||| : ||||| ||||| : ||||| | ||||| : 
Db   24 AELEFAQIIIIIIVVMVVIVCLLNHYKVSTRSFINRPNQSRREDGLPQEGCLWPSD 83
    ||||| ||||| : ||||| ||||| : ||||| : ||||| : 

QY   62 STVSGNGIPEPQVYAPRPTDRLAVPPFAQRERFHRFQPTYPYLQHEIDLPTISLDGE 121
    | | | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db   84 SAAPRLGASE--IMHAPRSRDRTAPSFIQRDRSFRQPTYPYVQHEIDLPTISLDGE 141
    | | | | : | | | | | | | | | | | | | | | | | | | | | | | |

QY   122 EPPYQGPGCTLQLRDPQQLELNRESVRAPPNRTIFDSDLMDSARL-GGPCPPSSNSGIS 180
    ||||| ||||| ||||| : ||||| ||||| ||||| : | ||||| ||||| |||||
Db   142 EPPYQGPGCTLQLRDPQQMELNRESVRAPPNRTIFDSLIDIAMYSGGPCPPSSNSGIS 201
    ||||| ||||| ||||| : ||||| ||||| ||||| : | ||||| ||||| |||||

QY   181 ATCYGSGRMGEGPPPTYSEVI GHYPGSSFHQHQSSGPSPILLEGTRLHHTHIAPLESAIW 240
    | : | ||||| ||||| : ||||| : ||||| : ||||| : ||||| : 
Db   202 ASTCSSNGRMGEGPPPTYSEVMGHHPGASFLHHQRS---NAHRGSRLQFQQ-NNAESTIVP 257
    | : | ||||| ||||| : ||||| : ||||| : ||||| : ||||| : 

QY   241 SKEKDQKGH 250
    | | | | : | : 
Db   258 IKGDKRKPGN 267
```

Search completed: December 4, 2003, 19:53:23
Job time : 32 secs

RESULT 15

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2003, 04:18:58 ; Search time 387 Seconds
(without alignments)
9111.986 Million cell updates/sec

Title: US-09-857-826B-44
Perfect score: 1061
Sequence: 1 tcctccttggttcgggtga.....cttgcaaaaaaaaaaaaaa 1061

Scoring table: OLIGO_NUC
Gapop 60.0 ; Gapext 60.0

Searched: 2201672 seqs, 1661799599 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	961	90.6	969	11 US-09-796-753-55	Sequence 55, Appl
2	926	87.3	1141	13 US-10-301-822-208	Sequence 208, App
3	926	87.3	1141	15 US-10-205-823-412	Sequence 412, App
4	926	87.3	1850	13 US-10-241-220-44	Sequence 44, Appl
5	926	87.3	4527	13 US-09-821-812-2	Sequence 2, Appli
6	912	86.0	1140	13 US-10-390-045-1	Sequence 1, Appli
7	902	85.0	4839	13 US-10-241-220-119	Sequence 119, App
8	902	85.0	4839	13 US-10-269-909-84	Sequence 84, Appl
9	902	85.0	4839	13 US-10-269-909-85	Sequence 85, Appl
10	891	84.0	1066	14 US-10-098-841-71	Sequence 71, Appl
11	788	74.3	806	13 US-10-241-220-45	Sequence 45, Appl
12	749	70.6	1321	10 US-09-934-249-1	Sequence 1, Appli
13	708	66.7	759	13 US-10-390-045-2	Sequence 2, Appli
14	701	66.1	861	10 US-09-934-249-3	Sequence 3, Appli
15	509	48.0	1583	15 US-10-000-256A-32	Sequence 32, Appl

C	16	313	29.5	693	10	US-09-934-249-14	Sequence 14, Appl
	17	269	25.4	467	11	US-09-918-995-2074	Sequence 2074, Ap
	18	102	9.6	368	10	US-09-783-590-3464	Sequence 3464, Ap
	19	60	5.7	60	13	US-09-908-975-13620	Sequence 13620, A
	20	50	4.7	65	10	US-09-783-590-3488	Sequence 3488, Ap
	21	44	4.1	878	10	US-09-934-249-12	Sequence 12, Appl
	22	44	4.1	1713	11	US-09-796-753-57	Sequence 57, Appl
	23	32	3.0	577	9	US-09-864-761-20542	Sequence 20542, A
	24	32	3.0	1964	9	US-09-864-761-3776	Sequence 3776, Ap
	25	29	2.7	401	9	US-09-864-761-3936	Sequence 3936, Ap
	26	29	2.7	446	9	US-09-864-761-20699	Sequence 20699, A
	27	29	2.7	475	10	US-09-934-249-15	Sequence 15, Appl
C	28	29	2.7	3444	13	US-10-293-582-16	Sequence 16, Appl
C	29	26	2.5	1493	13	US-10-029-386-25133	Sequence 25133, A
	30	26	2.5	7733	10	US-09-860-670-159	Sequence 159, App
	31	26	2.5	8429	9	US-09-738-885-3	Sequence 3, Appli
C	32	25	2.4	331	9	US-09-864-761-17053	Sequence 17053, A
C	33	25	2.4	465	9	US-09-864-761-230	Sequence 230, App
	34	25	2.4	470	9	US-09-864-761-3121	Sequence 3121, Ap
	35	25	2.4	522	9	US-09-864-761-19900	Sequence 19900, A
	36	25	2.4	522	13	US-10-029-386-22252	Sequence 22252, A
	37	25	2.4	571	13	US-10-029-386-9045	Sequence 9045, Ap
	38	25	2.4	756	13	US-10-029-386-20432	Sequence 20432, A
	39	24	2.3	24	13	US-10-390-045-5	Sequence 5, Appli
C	40	24	2.3	24	13	US-10-390-045-6	Sequence 6, Appli
C	41	24	2.3	24	13	US-10-390-045-10	Sequence 10, Appl
	42	24	2.3	58	15	US-10-106-698-3514	Sequence 3514, Ap
C	43	24	2.3	200	14	US-10-005-169-6	Sequence 6, Appli
C	44	24	2.3	569	13	US-10-029-386-3710	Sequence 3710, Ap
C	45	24	2.3	1884	9	US-09-853-386-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-09-796-753-55
; Sequence 55, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23


```

; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 55
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6)... (761)
US-09-796-753-55

```

	Query Match	90.6%;	Score 961;	DB 11;	Length 969;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 961;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	101	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGATGGTG	160		
DB	6	ATGGCGGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTGATGATGATGGTG	65		
QY	161	GTGGTGATCACGTGCCTGCTAGCCACTACAAGCTGTCTGCACGGTCTCTCATCAGCCGG	220		
DB	66	GTGGTGATCACGTGCCTGCTAGCCACTACAAGCTGTCTGCACGGTCTCTCATCAGCCGG	125		
QY	221	CACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCG	280		
DB	126	CACAGCCAGGGCGGAGGAGAGAGATGCCCTGTCTCTCAGAAGGATGCCTGTGGCCCTCG	185		
QY	281	GAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCCGCCCTCGGCC	340		
DB	186	GAGAGCACAGTGTCAAGCAACGGAAATCCAGAGCCGAGGTCTACGCCCCGCCCTCGGCC	245		
QY	341	ACCGACCGCCTGGCCGTGCGCCCTTCGCCCCAGCGGAGCGCTTCACCGCTTCCAGCCC	400		
DB	246	ACCGACCGCCTGGCCGTGCGCCCTTCGCCCCAGCGGAGCGCTTCACCGCTTCCAGCCC	305		
QY	401	ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTTCAGACGG	460		
DB	306	ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTTCAGACGG	365		
QY	461	GAGGAGCCCCACCCCTACCGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG	520		
DB	366	GAGGAGCCCCACCCCTACCGGGCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG	425		
QY	521	CTGGAACCTGAACCGGGAGTCGGTGGCGGCACCCCCAACAGAACCATCTTCGACAGTGAC	580		
DB	426	CTGGAACCTGAACCGGGAGTCGGTGGCGGCACCCCCAACAGAACCATCTTCGACAGTGAC	485		
QY	581	CTGATGGATAGTGCCAGGTGGGCGGCCCTGCCCCCCCAGCAGTAACTCGGGGCATCAGC	640		
DB	486	CTGATGGATAGTGCCAGGTGGGCGGCCCTGCCCCCCCAGCAGTAACTCGGGGCATCAGC	545		

RESULT 2

```

US-10-301-822-208
; Sequence 208, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(854)
US-10-301-822-208

```

Query Match	87.38;	Score 926;	DB 13;	Length 1141;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 976; Conservative	0;	Mismatches	1;	Indels

Qy 81 GTCTCTCGAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCTG 140
Db 76 GTCTCTCGAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCTG 135
Qy 141 TGGTGATGATGGTGTATGGTGGTGTATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTG 200
Db 136 TGGTGATGATGGTGTATGGTGGTGTATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTG 195
Qy 201 CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGAGAGAGATGCCCTGTCTCTAG 260
Db 196 CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGAGAGATGCCCTGTCTCTAG 255
Qy 261 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGAGG 320
Db 256 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGAGG 315
Qy 321 TCTACGCCCCGCTCGGCCCAACGACCGCCTGGCCGTGGCCCTTCGCCAGCGGGAGC 380
Db 316 TCTACGCCCCGCTCGGCCCAACGACCGCCTGGCCGTGGCCCTTCGCCAGCGGGAGC 375
Qy 381 GCTTCCACCGCTTCAGCCCAACCTATCCGTACCTGCAGCACAGATCGACCTGCCGCCA 440
Db 376 GCTTCCACCGCTTCAGCCCAACCTATCCGTACCTGCAGCACAGATCGACCTGCCGCCA 435
Qy 441 CCATCTCGCTGTAGACGGGAGGAGCCCAACCTACCGGCCCCCTGCACCTCCAGC 500
Db 436 CCATCTCGCTGTAGACGGGAGGAGCCCAACCTACCGGCCCCCTGCACCTCCAGC 495
Qy 501 TTCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGCACCCCAACA 560
Db 496 TTCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGCACCCCAACA 555
Qy 561 GAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGCGGCCCTGCCCCCA 620
Db 556 GAACCATCTTCGACAGTGACCTGATGGATAGTGCCAGGCTGGCGGCCCTGCCCCCA 615
Qy 621 GCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGCATGGAGGGCGC 680
Db 616 GCAGTAACCTCGGGCATCAGCGCCACGTGCTACGGCAGCGGGCGCATGGAGGGCGC 675
Qy 681 CGCCCACTACAGCGAGGTATCGGCCACTACCGGGGTCTCTCCAGCACAGCAGAGA 740
Db 676 CGCCCACTACAGCGAGGTATCGGCCACTACCGGGGTCTCTCCAGCACAGCAGAGA 735
Qy 741 GCAGTGGCGCCCTCTCTGTGGAGGGAGCCCGGCTCCACCAACACACATCGCGCCCC 800
Db 736 GCAGTGGCGCCCTCTCTGTGGAGGGAGCCCGGCTCCACCAACACACATCGCGCCCC 795
Qy 801 TAGAGAGCGAGCCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGG 860
Db 796 TAGAGAGCGAGCCCATCTGGAGCAAGAGAGGATAAACAGAAAGGACACCTCTCTAGG 855
Qy 861 GTCCCAAGGGGGCGGGCTGGGCTGCGTAGGTGAAAGGCAAGCACTCCGCGCTTCT 920
Db 856 GTCCCAAGGGGGCGGGCTGGGCTGCGTAGGTGAAAGGCAAGCACTCCGCGCTTCT 915
Qy 921 TAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCC 980
Db 916 TAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCC 975
Qy 981 CACCTCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGATGCACAAGCTAAGAGA 1040
Db 976 CACCTCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGATGCACAAGCTAAGAGA 1035
Qy 1041 GCTTGCAAAAAA 1057
Db 1036 GCTTGCAAAAAA 1052

RESULT 3
US-10-205-823-412
; Sequence 412, Application US/10205823

; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glat, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 412
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-412

Query Match 87.3%; Score 926; DB 15; Length 1141;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 81 GTCTCTCGAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCTG 140
Db 76 GTCTCTCGAAACAGGCAATGGCGAGCTGGAGTTTGTTCAGATCATCATCTG 135
Qy 141 TGGTGATGATGGTGTATGGTGGTGTATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTG 200
Db 136 TGGTGATGATGGTGTATGGTGGTGTATCACGTGCCCTGCTGAGCCACTACAAGCTGTCTG 195
Qy 201 CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGAGAGATGCCCTGTCTCTAG 260
Db 196 CACGGTCCTTCATCAGCCGGCACAGCCAGGGCGGAGAGAGATGCCCTGTCTCTAG 255
Qy 261 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGAGG 320
Db 256 AAGGATGCCTGTGGCCCTCGGAGAGCACAGTGTCAAGCAACGGAATCCCAGAGCCGAGG 315
Qy 321 TCTACGCCCCGCTCGGCCCAACCGCCTGGCCGTGGCCCTTCGCCAGCGGGAGC 380
Db 316 TCTACGCCCCGCTCGGCCCAACCGCCTGGCCGTGGCCCTTCGCCAGCGGGAGC 375
Qy 381 GCTTCCACCGCTTCAGCCCAACCTATCCGTACCTGCAGCACAGATCGACCTGCCGCCA 440
Db 376 GCTTCCACCGCTTCAGCCCAACCTATCCGTACCTGCAGCACAGATCGACCTGCCGCCA 435
Qy 441 CCATCTCGCTGTAGACGGGAGGAGCCCAACCTACCGTACCGAGGGCCCTGCACCTCCAGC 500
Db 436 CCATCTCGCTGTAGACGGGAGGAGCCCAACCTACCGTACCGAGGGCCCTGCACCTCCAGC 495
Qy 501 TTCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGCACCCCAACA 560
Db 496 TTCGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTGGCGCACCCCAACA 555

QY	561	GAACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGCGGCCCTGCCCCC	620
DB	556	GAACCATCTTCGACAGTACCTGATGGATAGTGCAGGCTGGCGGCCCTGCCCCC	615
QY	621	GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCG	680
DB	616	GCAGTAACTCGGGCATCAGCGCCACGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCG	675
QY	681	CGCCACCTACAGCGAGGTCACTCGGCCACTACCGGGGTCTCTCTCCAGCACAGCAGA	740
DB	676	CGCCACCTACAGCGAGGTCACTCGGCCACTACCGGGGTCTCTCTCCAGCACAGCAGA	735
QY	741	GCAGTGGCGCGCCCTCCTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCC	800
DB	736	GCAGTGGCGCGCCCTCCTTGCTGGAGGGGACCCGGCTCCACCACACACATCGCGCCCC	795
QY	801	TAGAGAGCGCAGCCATCTGGAGCAAAAGAGAAGGATAAACAGAAAGGACACCTCTCTAGG	860
DB	796	TAGAGAGCGCAGCCATCTGGAGCAAAAGAGAAGGATAAACAGAAAGGACACCTCTCTAGG	855
QY	861	GTCCCCAGGGGGCCGGCTGGGGCTGCGTAGTGAAAGGCAGAACACTCCGCGCTTCT	920
DB	856	GTCCCCAGGGGGCCGGCTGGGGCTGCGTAGTGAAAGGCAGAACACTCCGCGCTTCT	915
QY	921	TAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCC	980
DB	916	TAGAAGAGGAGTGAGAGGAAGCGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCC	975
QY	981	CACCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGA	1040
DB	976	CACCTCCCTGTGTATAAATATTTACATGTGATGTCTGGTCTGAATGCACAAGCTAAGAGA	1035
QY	1041	GCTTGCAAAAAA	1057
DB	1036	GCTTGCAAAAAA	1052

RESULT 4

RESULT 4
 US-10-241-220-44
 ; Sequence 44, Application US/10241220
 ; Publication No. US20030148408A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Frantz, Gretchen
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Phillips, Heidi
 ; APPLICANT: Polakis, Paul
 ; APPLICANT: Spencer, Susan
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wu, Thomas
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
 ; TREATMENT OF TUMOR
 ; FILE REFERENCE: P5010R1-US
 ; CURRENT APPLICATION NUMBER: US/10/241,220
 ; CURRENT FILING DATE: 2002-12-13
 ; NUMBER OF SEQ ID NOS: 120
 ; SEQ ID NO 44
 ; LENGTH: 1850
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-241-220-44

	Query Match	87.3%	Score 926;	DB 13;	Length 1850;
	Best Local Similarity	99.9%;	Pred. No. 0;		
	Matches 976;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	81	GTCTCCTCGGAAACCAAGCAATGGCGGAGCTGGAGTTTGTTTCAGATCATCATCATCGTGG	140		
Dd	76	GTCTCCTCGGAAACCAAGCAATGGCGGAGCTGGAGTTTGTTTCAGATCATCATCATCGTGG	135		
QY	141	TGGTGATGATGGTGATGGTGGTGATCACGTGCCGTGCTGAGCCACTACAAGCTGTCTG	200		
Dd	136	TGGTGATGATGGTGATGGTGGTGATCACGTGCCGTGCTGAGCCACTACAAGCTGTCTG	195		

RESULT 5
US-09-821-812-2
; Sequence 2, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; TITLE OF INVENTION: Nucleic Acids
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28

RESULT 5

RESOLUTION 3
US-09-821-812-2
; Sequence 2, Application US/09821812
; Publication No. US20030166520A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Prostate Specific
; TITLE OF INVENTION: Nucleic Acids
; FILE REFERENCE: P-IS 4373
; CURRENT APPLICATION NUMBER: US/09/821,812
; CURRENT FILING DATE: 2001-03-28

Matches 976; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 6

Query Match 86.0%; Score 912; DB 13; Length 1140;

Query Match 86.0%; Score 912; DB 13; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	95	CAGGCAATGGCGAGCTGGAGTTGTTT	CAGATCATCATCATCGTGGTGATGGTG	154
Db	89	CAGGCAATGGCGAGCTGGAGTTGTTT	CAGATCATCATCATCGTGGTGATGGTG	148
QY	155	ATGGTGGTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCTTTCATC	214	
Db	149	ATGGTGGTGGTGATCACGTGCTGCTGAGCCACTACAAGCTGTCTGCACGGTCCTTTCATC	208	
QY	215	AGCCGGCACAGCCAGGGGCGGAGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGG	274	
Db	209	AGCCGGCACAGCCAGGGGCGGAGGAGAGAAGATGCCCTGTCTCTCAGAAGGATGCCTGTGG	268	
QY	275	CCCTCGGAGAGCACAGTGTTCAGGCAACGGAATCCCAGAGCCGCGAGGTCTACGCCCCCGCCT	334	
Db	269	CCCTCGGAGAGCACAGTGTTCAGGCAACGGAATCCCAGAGCCGCGAGGTCTACGCCCCCGCCT	328	
QY	335	CGGCCCAACCGACCGCCTGGCGTGC CGCCCTTTCGCCACGCGGAGCGCTTCCACCGCTTC	394	
Db	329	CGGCCCAACCGACCGCCTGGCGTGC CGCCCTTTCGCCACGCGGAGCGCTTCCACCGCTTC	388	
QY	395	CAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCCATCTCGCTGTCA	454	
Db	389	CAGCCCCACCTATCCGTACCTGCAGCACGAGATCGACCTGCCACCCATCTCGCTGTCA	448	
QY	455	GACGGGGAGGAGCCCCACCTACACAGGGCCCCCTGCACCCCTCCAGTTTCGGGACCCCCGAG	514	
Db	449	GACGGGGAGGAGCCCCACCTACACAGGGCCCCCTGCACCCCTCCAGTTTCGGGACCCCCGAG	508	


```
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-84

Query Match      85.0%; Score 902; DB 13; Length 4839;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 105 CGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 164
    |||||
Db 430 CGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 489

Qy 165 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGGCACA 224
    |||||
Db 490 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGGCACA 549

Qy 225 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 284
    |||||
Db 550 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 609

Qy 285 GCACAGTGTGAGGCAACGGAAATCCAGAGCCGCGAGTCTACGCCCGCCTCGGACCG 344
    |||||
Db 610 GCACAGTGTGAGGCAACGGAAATCCAGAGCCGCGAGTCTACGCCCGCCTCGGACCG 669

Qy 345 ACCGCTGGCCGTGCCCGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 404
    |||||
Db 670 ACCGCTGGCCGTGCCCGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 729

Qy 405 ATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGCTGTGAGCGGGAGG 464
    |||||
Db 730 ATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGCTGTGAGCGGGAGG 789

Qy 465 AGCCCCACCCCTACCGGGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGG 524
    |||||
Db 790 AGCCCCACCCCTACCGGGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGG 849

Qy 525 AACTGAACCGGGAGTCGGTGCGGCAACAGAAACAGAAACAGAAACAGAAACAGAAACAG 584
    |||||
Db 850 AACTGAACCGGGAGTCGGTGCGGCAACAGAAACAGAAACAGAAACAGAAACAGAAACAG 909

Qy 585 TGGATAGTCCAGGTCGGGCGGCGCATGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 644
    |||||
Db 910 TGGATAGTCCAGGTCGGGCGGCGCATGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 969

Qy 645 CGTGCTACGGCAGCGGGCGGCGCATGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 704
    |||||
Db 970 CGTGCTACGGCAGCGGGCGGCGCATGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1029

Qy 705 GCCACTACCCGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 764
    |||||
Db 1030 GCCACTACCCGGGTCTCTCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1089

Qy 765 AGGGAGCCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 824
    |||||
Db 1090 AGGGAGCCCGGCTCCACCACACACATCGCGCCCTTAGAGAGCGCAGCCATCTGGAGCA 1149

Qy 825 AAGAGAAGGATAAACAAGAAAGGACACCCCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 884
    |||||
Db 1150 AAGAGAAGGATAAACAAGAAAGGACACCCCTCTCTAGGGTCCCGAGGGGGCGGGCTGGGG 1209

Qy 885 CTGCTAGGTGAAAGGCGAAGCACTCCCGGCTTCTTAGAAGAGGAGTGAGAGGAAGCGG 944
    |||||
```

```
Db 1210 CTGCGTAGGTGAAAGGCGAGAACACTCCGGCGCTTCTTAGAAGAGGAGTGAGAGGAGCG 1269

Qy 945 GGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTTA 1004
    |||||
Db 1270 GGGGGCGCAGCAACGCATCGTGTGGCCCTCCCTCCACCTCCCTGTGTATATAATATTTA 1329

Qy 1005 CATGTGATGTCTGCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1057
    |||||
Db 1330 CATGTGATGTCTGCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAA 1382

RESULT 9
US-10-269-909-85
; Sequence 85, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUZIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 4839
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-269-909-85
```

```
Query Match      85.0%; Score 902; DB 13; Length 4839;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 952; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 105 CGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 164
    |||||
Db 430 CGGAGCTGGAGTTGTTTCAGATCATCATCATCGTGGTGGTGGTGGTGGTGG 489

Qy 165 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGGCACA 224
    |||||
Db 490 TGATCAGCTGCTGTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGGCACA 549

Qy 225 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 284
    |||||
Db 550 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 609

Qy 285 GCACAGTGTGAGGCAACGGAAATCCAGAGCCGCGAGTCTACGCCCGCCTCGGACCG 344
    |||||
Db 610 GCACAGTGTGAGGCAACGGAAATCCAGAGCCGCGAGTCTACGCCCGCCTCGGACCG 669

Qy 345 ACCGCTGGCCGTGCCCGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 404
    |||||
Db 670 ACCGCTGGCCGTGCCCGCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 729

Qy 405 ATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGCTGTGAGCGGGAGG 464
    |||||
Db 730 ATCCGTACCTGCAGCAGCAGATCGACCTGCCGCCACCATCTCGCTGTGAGCGGGAGG 789

Qy 465 AGCCCCACCCCTACCGGGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGG 524
    |||||
Db 790 AGCCCCACCCCTACCGGGCCCTGACCCCTCCAGCTTCGGGACCCCGAGCAGCTGG 849

Qy 525 AACTGAACCGGGAGTCGGTGCGGCAACAGAAACAGAAACAGAAACAGAAACAGAAACAG 584
    |||||
Db 850 AACTGAACCGGGAGTCGGTGCGGCAACAGAAACAGAAACAGAAACAGAAACAGAAACAG 909

Qy 585 TGGATAGTCCAGGTCGGGCGGCGCATGGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 644
    |||||
```


D _b	910	TGGATAGTCCAGGCTGGCGGCCCTTGCCCCCAGCAGTAACTCGGGCATCAGCGCCA	969
Q _y	645	CGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCGCCCACCTACAGCGAGGTCAATCG	704
D _b	970	CGTGCTACGGCAGCGCGGGCGCATGGAGGGCGCGCGCCCACCTACAGCGAGGTCAATCG	1029
Q _y	705	GCCACTACCCGGGTCCTCTTCAGCACACCAGCAGAGCAGTGGCCGCCCTCCTTGCTGG	764
D _b	1030	GCCACTACCCGGGTCCTCTTCAGCACACCAGCAGAGCAGTGGCCGCCCTCCTTGCTGG	1089
Q _y	765	AGGGGACCCGGCTCCACCAACAACAATCGCGCCCCCTAGAGAGCGCAGCCATCTGGAGCA	824
D _b	1090	AGGGGACCCGGCTCCACCAACAACAATCGCGCCCCCTAGAGAGCGCAGCCATCTGGAGCA	1149
Q _y	825	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGTCCCCCAGGGGGCCGGGCTGGGG	884
D _b	1150	AAGAGAAGGATAAACAGAAAGGACACCTCTCTAGGTCCCCCAGGGGGCCGGGCTGGGG	1209
Q _y	885	CTGCGTAGGTGAAGGAGCAGAACACTCCGCGCTTTAGAAAGAGGAGTGAGAGGAAGCG	944
D _b	1210	CTGCGTAGGTGAAGGAGCAGAACACTCCGCGCTTTAGAAAGAGGAGTGAGAGGAAGCG	1269
Q _y	945	GGGGGGCAGCAAACGCATCGTGTGGCCCTCCCCCTCCACCTCCCTGTGTATAAATATTTA	1004
D _b	1270	GGGGGGCAGCAAACGCATCGTGTGGCCCTCCCCCTCCACCTCCCTGTGTATAAATATTTA	1329
Q _y	1005	CATGTGATGTCCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA	1057
D _b	1330	CATGTGATGTCCTGGTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAA	1382

RESULT 10

```

US-10-098-841-71
; Sequence 71, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 71
; LENGTH: 1066
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..(867)

```


Db 942 AACTGAACGGGAGTCGGTGGCGCACCCCAAAACAGAACCATCTTCGACAGTGACCTGA 1001
Qy 585 TGGATAGTCCAGGCTGGCGGCCCTGCCCGCCAGCAGTAACCTCGGGCATCAGCGCA 644
Db 1002 TGGATAGTCCAGGCTGGCGGCCCTGCCCGCCAGCAGTAACCTCGGGCATCAGCGCA 1061
Qy 645 CGTGCTACGGCAGCGCGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTCTCG 704
Db 1062 CGTGCTACGGCAGCGCGCGGCATGGAGGGGCGCGCCACCTACAGCGAGGTCTCG 1121
Qy 705 GCCACTACCGGGGTCTCTTCCAGCACAGCAGAGTGGGCGCCCTCTTGTGG 764
Db 1122 GCCACTACCGGGGTCTCTTCCAGCACAGCAGAGTGGGCGCCCTCTTGTGG 1181
Qy 765 AGGGACCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA 824
Db 1182 AGGGACCGGCTCCACCACACACATCGCGCCCTAGAGAGCGCAGCCATCTGGAGCA 1241
Qy 825 AAGAGAGGATAAAACAGAAAGGACACCCCTCTTAGGGTCCCGAGGGGCGGGCTGGG 884
Db 1242 AAGAGAGGATAAAACAGAAAGGACACCCCTCTTAGGGTCCCGAGGGGCGGGCTGGG 1301
Qy 885 CTGCGTAGTGAAAGGCAG 904
Db 1302 CTGCGTAGTGAAAGGCAG 1321

RESULT 13
US-10-390-045-2
; Sequence 2, Application US/10390045
; Publication No. US20030170713A1
; GENERAL INFORMATION:
; APPLICANT: SRIVASTAVA, SHIV
; APPLICANT: MOUL, JUDD W.
; APPLICANT: XU, LINDA L.
; APPLICANT: SEGAWA, TAKEHIKO
; TITLE OF INVENTION: PROSTATE-SPECIFIC ANDROGEN-SIGNALING-ASSOCIATED
; TITLE OF INVENTION: POYNUCLEOTIDE ARRAY
; FILE REFERENCE: 04995.0057-0000
; CURRENT APPLICATION NUMBER: US/10/390,045
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US/09/769,482
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,772
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,045
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-390-045-2

Query Match 66.7%; Score 708; DB 13; Length 759;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 101 ATGGCGGAGTGGAGTTTGTTCAGATCATCATCGTGGTGGTATGATGGTATGGT 160
Db 1 ATGGCGGAGTGGAGTTTGTTCAGATCATCATCGTGGTGGTATGATGGTATGGT 60
Qy 161 GTGGTGATCAGTGCCTGTCTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGG 220
Db 61 GTGGTGATCAGTGCCTGTCTGAGCCACTACAAGCTGTCTGACGGTCTTCATCAGCCGG 120
Qy 221 CACAGCCAGGGCGGAGGAGAGAAGATGCGCTGTCTCCTCAGAAGGATGCTGTGGCCCTCG 280
Db 121 CACAGCCAGGGCGGAGGAGAGAAGATGCGCTGTCTCCTCAGAAGGATGCTGTGGCCCTCG 180
Qy 281 GAGAGCACAGTGTCTAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCCTCGGCC 340

Db 181 GAGAGCACAGTGTCTAGGCAACGGAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCC 240
Qy 341 ACCGACCGCCTGGCGCGCCCTTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGGCC 400
Db 241 ACCGACCGCCTGGCGCGCCCTTTCGCCCCAGCGGAGCGCTTCCACCGCTTCCAGGCC 300
Qy 401 ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTCTCAGACGG 460
Db 301 ACCTATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGTGTCTCAGACGG 360
Qy 461 GAGGAGCCCCCACCCCTACAGGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 520
Db 361 GAGGAGCCCCCACCCCTACAGGGGCCCCCTGCACCCCTCCAGCTTCGGGACCCCGAGCAGCAG 420
Qy 521 CTGGAATCTGAACCGGGAGTGGTGGCGGCACCCCCCAACAGAACCATCTTCGACAGTGAC 580
Db 421 CTGGAATCTGAACCGGGAGTGGTGGCGGCACCCCCCAACAGAACCATCTTCGACAGTGAC 480
Qy 581 CTGATGGATAGTCCAGGCTGGGCGGCCCTTGCCCCCAGCAGTAACCTCGGGCATCAGC 640
Db 481 CTGATGGATAGTCCAGGCTGGGCGGCCCTTGCCCCCAGCAGTAACCTCGGGCATCAGC 540
Qy 641 GCCACGTGCTACGGCAGCGCGGGCGGCATGGAGGGGCGGCCCCACCTACAGCGAGGTC 700
Db 541 GCCACGTGCTACGGCAGCGCGGGCGGCATGGAGGGGCGGCCCCACCTACAGCGAGGTC 600
Qy 701 ATCGGCCACTACCGGGGTCTCTCTCCAGCACAGCAGCAGTGGGCGCCCTCTCTT 760
Db 601 ATCGGCCACTACCGGGGTCTCTCTCCAGCACAGCAGCAGTGGGCGCCCTCTCTT 660
Qy 761 CTGAGGGGACCCGGTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGG 820
Db 661 CTGAGGGGACCCGGTCCACCACACACATCGCGGCCCTTAGAGAGCGCAGCCATCTGG 720
Qy 821 AGCAAAAGAGAAGGATAAAACAGAAAGGACACCCCTCTCTAG 859
Db 721 AGCAAAAGAGAAGGATAAAACAGAAAGGACACCCCTCTCTAG 759

RESULT 14
US-09-934-249-3
; Sequence 3, Application US/09934249
; Patent No. US20020115081A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Richard T.
; APPLICANT: Landschulz, Katherine T.
; APPLICANT: Turi, Thomas G.
; APPLICANT: Thompson, John F.
; APPLICANT: Kennedy, Scott P.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR CONDITIONS
; FILE REFERENCE: P0738/7001/ERP/KA
; CURRENT APPLICATION NUMBER: US/09/934,249
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/227,159
; PRIOR FILING DATE: 2000-08-22
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 861
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)... (861)
US-09-934-249-3

Query Match 66.1%; Score 701; DB 10; Length 861;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 751; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 105 CGAGCTGGAGTTTGTTCAGATCATCATCGTGGTGGTATGATGGTGGTGG 164

Db 110 CGGAGCTGGAGTTTGTTCAGATCATCATCATCGTGGTGGTGATGATGGTGTGG 169

Qy 165 TGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA 224

Db 170 TGATCAGCTGCCTGCTGAGCCACTACAAGCTGTCTGCACGGTCTTCATCAGCCGGCACA 229

Qy 225 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 284

Db 230 GCCAGGGCGGAGGAGAGAGATGCCCTGTCTCAGAAAGGATGCCTGTGGCCCTCGGAGA 289

Qy 285 GCACAGTGTAGGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCCACCG 344

Db 290 GCACAGTGTAGGCAACGGAAATCCAGAGCCGAGGTCTACGCCCGCCTCGGCCACCG 349

Qy 345 ACCGCTGGCCGCTGCCGCCCTTCGCCCCAGCGGAGCGCTTCACCGCTTCAGCCCACT 404

Db 350 ACCGCTGGCCGCTGCCGCCCTTCGCCCCAGCGGAGCGCTTCACCGCTTCAGCCCACT 409

Qy 405 ATCCGTACCTGCAGCACGAGATCGACCTGCCGCCACCATCTCGCTGTAGACGGGAGG 464

Db 410 ATCCGTACCTGCAGCACGAGATCGACCTGCCACCACCATCTCGCTGTAGACGGGAGG 469

Qy 465 AGCCCCACCTACACAGGGCCCTGACACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 524

Db 470 AGCCCCACCTACACAGGGCCCTGACACCTCCAGCTTCGGGACCCCGAGCAGCAGCTGG 529

Qy 525 AACTGAACCGGAGTCGGTGGCGGCCCTGCCGCCAGCAAGTAACCTCGGGCATCAGCGCCA 644

Db 530 AACTGAACCGGAGTCGGTGGCGGCCCTGCCGCCAGCAAGTAACCTCGGGCATCAGCGCCA 649

Qy 585 TGGATAGTCCAGGCTGGCGGCCCTGCCGCCAGCAAGTAACCTCGGGCATCAGCGCCA 644

Db 590 TGGATAGTCCAGGCTGGCGGCCCTGCCGCCAGCAAGTAACCTCGGGCATCAGCGCCA 649

Qy 645 CGTGTACGGCAGCGCGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCTATCG 704

Db 650 CGTGTACGGCAGCGCGGCGCATGGAGGGCGCGCCGCCACCTACAGCGAGGTCTATCG 709

Qy 705 GCCACTACCCGGGTCTCTTCAGACACCAGCAGAGCAGTGGGCCGCCCTCTCTTGTGG 764

Db 710 GCCACTACCCGGGTCTCTTCAGACACCAGCAGAGCAGTGGGCCGCCCTCTCTTGTGG 769

Qy 765 AGGGACCCGGCTCCACCACACACATCGGCCCTTAGAGCGCAGCCATCTGGAGCA 824

Db 770 AGGGACCCGGCTCCACCACACACATCGGCCCTTAGAGCGCAGCCATCTGGAGCA 829

Qy 825 AAGAGAAGGATAAACAGAAAGGACACCTCTC 856

Db 830 AAGAGAAGGATAAACAGAAAGGACACCTCTC 861

RESULT 15

US-10-000-256A-32

; Sequence 32, Application US/10000256A

; Publication No. US20030039983A1

; GENERAL INFORMATION:

; APPLICANT: Sun, Yongming

; APPLICANT: Recipon, Herve

; APPLICANT: Chen, Sei-Yu

; APPLICANT: Liu, Chenghua

; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and

; FILE REFERENCE: DEX-0259

; CURRENT APPLICATION NUMBER: US/10/000,256A

; CURRENT FILING DATE: 2001-11-01

; PRIOR APPLICATION NUMBER: 60/244,782

; PRIOR FILING DATE: 2000-11-01

; NUMBER OF SEQ ID NOS: 240

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 32

; LENGTH: 1583

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-000-256A-32

Query Match 48.0%; Score 509; DB 15; Length 1583;

Best Local Similarity 99.8%; Pred. No. 6.3e-228;

Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 313 GCGCAGGTCTACGCCCGCTCGGCCACCGACCGCTTGGCGCTTGGCCCTTCGCCCA 372

Db 827 GCGCAGGTCTACGCCCGCTCGGCCACCGACCGCTTGGCGCTTGGCCCTTCGCCCA 886

Qy 373 GCGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGACCT 432

Db 887 GCGGAGCGCTTCCACCGCTTCCAGCCCACTATCCGTACCTGCAGCACGAGATCGACCT 946

Qy 433 GCGGCCACCATCTCGCTGTAGACGGGAGGAGCCCCACCTACAGGGCCCCCTGCAC 492

Db 947 GCCACCCACCATCTCGCTGTAGACGGGAGGAGCCCCACCTACAGGGCCCCCTGCAC 1006

Qy 493 CCTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTCGCGCAC 552

Db 1007 CCTCAGCTTCGGGACCCCGAGCAGCAGCTGGAACCTGAACCGGGAGTCGGTCGCGCAC 1066

Qy 553 CCAAAACAGAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGCTGGCGGCCCTG 612

Db 1067 CCAAAACAGAACCATCTTCGACAGTGACCTGATGATAGTGCCAGGCTGGCGGCCCTG 1126

Qy 613 CCCCCCAGCAGTAACCTCGGCGATCAGCGCCACGTGCTACGGCAGCGCGGCGCATGGA 672

Db 1127 CCCCCCAGCAGTAACCTCGGCGATCAGCGCCACGTGCTACGGCAGCGCGGCGCATGGA 1186

Qy 673 GGGCGCGCGCCCACTACAGCAGGTCATCGGCCACTACCCGGGGTCTCTTCCAGCA 732

Db 1187 GGGCGCGCGCCCACTACAGCAGGTCATCGGCCACTACCCGGGGTCTCTTCCAGCA 1246

Qy 733 CCAGCAGCAGTGGGCGCCCTCTTGTGGAGGGGACCCGGCTCCACCACACACAT 792

Db 1247 CCAGCAGCAGTGGGCGCCCTCTTGTGGAGGGGACCCGGCTCCACCACACACAT 1306

Qy 793 CGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAGAGGATAAAACAGAAAGGACACC 852

Db 1307 CGCGCCCTTAGAGAGCGCAGCCATCTGGAGCAAAGAGAGGATAAAACAGAAAGGACACC 1366

Qy 853 TCTCTAGGGTCCCCAGGGG 872

Db 1367 TCTCTAGGGTCCCCAGGGG 1386

Search completed: December 8, 2003, 06:25:42

Job time : 391 secs

THIS PAGE BLANK (USPTO)